MASTER'S DEGREE THESIS

Master of Science in Advanced Mathematics and Mathematical Engineering

Title: Discrete duality principle in different random graph models

Author: Thomas Lesgourgues

Advisor: Marc Noy & Guillem Perarnau

Department: Department of Mathematics

Academic year: 2018-2019

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya
Facultat de Matemàtiques i Estadística

Master in Advanced Mathematics and Mathematical Engineering

Master's thesis

Discrete Duality Principle in different random graph models

Thomas Lesgourgues

Supervised by Marc Noy & Guillem Perarnau

June 25, 2019

Acknowledgements

I would like to express my gratitude to my supervisors Marc Noy and Guillem Perarnau for their availability and engagement through this master thesis. This has been a learning process full of pleasant discussions and discoveries. Together with some other professors (such as Oriol Serra and Juanjo Rué), thanks to their will to develop and implement a culture of collaboration into their research groups, I have been able to discover and immerse myself fully into a potential new career. In the hope of possible future collaborations, my sincere acknowledgements.

More than anyone, I am forever grateful to Hélène for her constant support, love and guidance, always with me in whatever crazy idea I pursue.

Abstract

For a given random graph, a connected component that contains a finite fraction of the entire graph's vertices is called giant. The study of these components started with the Erdős-Rényi model, where it has been proven that removing the (unique) giant component from a random graph is essentially equivalent to another random graph in the same model with different known parameters. This is called the discrete duality principle. In this report we aim at presenting this principle in its historical Erdős-Rényi settings, and then to present more recent generalisations made on random graphs with given degree sequences.

Keywords

Random graph, Giant Component, Discrete Duality Principle, Branching process, Erdős-Rényi model, Configuration model, Switching.

Contents

1. Introduction & overview

Perhaps the most studied phenomenon in the field of random graphs is the behaviour of the size of the largest component in $G(n,p)$* when $p = c/n$. For $c < 1$ the size of the largest component is almost surely† $O(\log n)$. For $c = 1$ the size of the largest component is a.s. $\Theta(n^{2/3})$. And for $c > 1$ a.s. the size of the largest component is $\Theta(n)$ while the size of the second largest component is $O(\log n)$ (see [AS92]). For $c > 1$, this largest component is commonly referred to as the *giant component* and the point $p = 1/n$ is referred to as the *critical point* or the *double jump threshold*.

Many results have been reached from this point, with notably a Central Limit Theorem for the size of this giant component. Another very famous result is the *Discrete Duality Principle* : Given a graph $G \in G(n,p)$, with $p = c/n$ and $c > 1$, the graph obtained by removing from G its giant component is essentially equivalent to a graph on $G(n',p')$ for specific known n' and p'.

However this traditional Erdős-Rényi model is of little use in modelling the type of complex networks which researchers study nowadays. Modern networks are of diverse nature and usually exhibit inhomogeneity among their nodes and correlations among their edges. This motivates the study, for a fixed degree sequence $\mathcal{D} = (d_1, \dots, d_n)$, of graphs $G(\mathcal{D})$ on n vertices where vertex i has degree d_i. In [MR95], Molloy and Reed showed that a similar phenomenon occurs among random graphs with a fixed degree sequence, provided that the degree sequence under consideration satisfied certain technical conditions. Essentially, they considered random graphs on n vertices with $\lambda_i n + o(n)$ vertices of degree i, for some fixed sequence $\lambda_0, \lambda_1, \dots$. They introduced the parameter $\mathcal{Q} = \sum_{i \geq 0} i(i-2)\lambda_i$ and showed that, if $\mathcal{Q} < 0$, then a.s. the size of the largest component is $O(\omega^2 \log n)$, where ω is the highest degree in the graph, and if $Q > 0$, then a.s. the size of the largest component is $\Theta(n)$, and the size of the second largest component is $O(\log n)$. In a subsequent article, they derive a discrete duality principle for these degree sequences, proving that Erdős-Rényi result is a particular case of their model.

This work has attracted considerable attention and has been applied to random models of a wide range of complex networks such as the World Wide Web or biological systems. Furthermore, many authors have obtained related results which formalize the Molloy-Reed heuristic argument under different sets of technical conditions, among many others, see Kang and Seierstad [KS08], Pittel [Pit08], or Riordan [Rio12] for example.

One of the main technical conditions under which these results on the existence of a giant component hold, is that the vertices of high degree do not have a large impact on the structure of the graph. However, in many real-life networks‡, the vertices of high degree (called hubs) have a crucial role in several of the network properties. Hence, often these results cannot be directly applied to real world networks where hubs are present, and ad-hoc approaches are needed for each particular network (see for instance [ACL00]).

To remedy this issue, in [JPRR18], Joos, Perarnau, Rautenbach and Reed extended these results to any degree sequence. They characterize the condition upon which a random graph with given degree sequence has a giant component. They only require to avoid the case where almost all graphs have degree 2. Besides the fact that it is a relatively minor technical condition, they also show that if it is not satisfied, both the probability that $G(\mathcal{D})$ has a giant component and the probability that $G(\mathcal{D})$ has no giant component are bounded away from 1.

They showed that their results are generalizations of the existence of the giant component in the settings

*$G(n,p)$ is the random graph with n vertices where each edge appears independently with probability p.

†See appendix A.1 for asymptotic notations

‡In general, we recommend [Hof16] for a general presentation of real-world networks and their properties.

from Molloy and Reed (and therefore for the Erdős-Rényi model). However if the existence of the giant component is now perfectly solved, it remains to show the conditions under which this component is unique (or not), and to deduce a Discrete Duality Principle.

2. Branching Process

Branching processes will be used extensively throughout these notes, describing the connected components of random graphs in the Erdős-Rényi model. Special attention will be given to Poisson and Binomial branching processes.

2.1 Galton-Watson branching process

A branching process is a model for population evolving in time. It starts with a single root node, and then each particle (a.k.a nodes) independently gives birth to a number of children (a.k.a offspring) with the same probability distribution. We denote this distribution by $\{p_i\}_{i=0}^{\infty}$ with

$$p_i = \mathbb{P}[\text{individual has } i \text{ children}]$$

Let X be the offspring distribution, and $\{X_{n,i}\}_{n,i\geq 1}$ a double infinite array of i.i.d random variable. $X_{n,i}$ denotes the number of children generated at time n by the parent i, so that for all n, i, $X_{n,i} \sim X$. We denote by Z_n the number of individuals in the n^{th} generation. Then Z_n satisfies the recursion :

$$\begin{cases} Z_0 = 1 \\ Z_n = \sum_{i=1}^{Z_{n-1}} X_{n,i} \end{cases}$$

The main result detailed below (see Theorem 2.1) is that when $\mathbb{E}[X] \leq 1$ the population dies out with probability 1, while if $\mathbb{E}[X] > 1$, there is a non-zero probability that the population will not become extinct. We denote the *extinction probability* by

$$\eta = \mathbb{P}[\exists n,\ Z_n = 0] \tag{1}$$

We recall the definition of G_X, the *probability generating function* of a random variable X :

$$G_X(s) = \mathbb{E}[s^X]$$

If X is a discrete random variable, taking value k with probability p_k, then

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k\geq 0} p_k s^k$$

Theorem 2.1 (Survival v.s. extinction)**.** *For a branching process with i.i.d offspring X, $\eta = 1$ when $\mathbb{E}[X] < 1$, while $\eta < 1$ when $\mathbb{E}[X] > 1$. When $\mathbb{E}[X] = 1$, and $\mathbb{P}[X = 1] < 1$, then $\eta = 1$. Moreover the extinction probability η is the smallest solution in $[0, 1]$ of*

$$\eta = G_X(\eta) \tag{2}$$

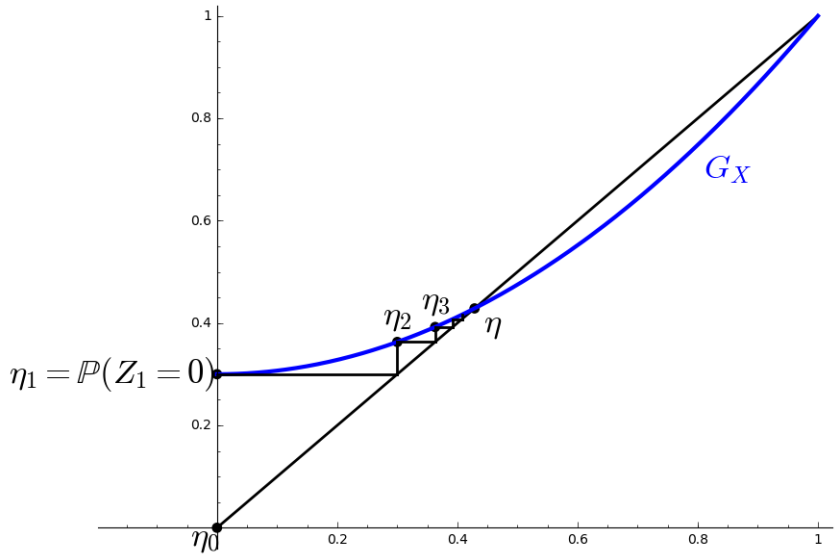


Figure 1: $\lim_{n\to\infty} \eta_n = \eta$

Proof. Let $G_n(s) = \mathbb{E}[s^{Z_n}]$. Conditioning on the first generation:

$$G_n(s) = \sum_i p_i \mathbb{E}[s^{Z_n} \mid Z_1 = i]$$

If $Z_1 = i$ then each nodes $k = 1, \dots, i$ has for n^{th} generation Z_{n-1}, all i.i.d.

$$\mathbb{E}[s^{Z_n} \mid Z_1 = i] = \mathbb{E}[s^{i \cdot Z_{n-1}}] = \mathbb{E}[s^{Z_{n-1}}]^i = G_{n-1}(s)^i$$

So that

$$G_n(s) = \sum_i p_i \mathbb{E}[s^{Z_n} \mid Z_1 = i] = \sum_i p_i G_{n-1}(s)^i$$

$$G_n(s) = G_X(G_{n-1}(s)) \tag{3}$$

For any integer valued random variable Y, $\mathbb{P}[Y = 0] = G_Y(0)$, therefore, taking $s = 0$,

$$\mathbb{P}[Z_n = 0] = G_X(\mathbb{P}[Z_{n-1} = 0])$$

Writing the probability of extinction of each step $\eta_n = \mathbb{P}[Z_n = 0]$,

$$\eta_n = G_X(\eta_{n-1}) \tag{4}$$

Using the fact that $Z_n = 0$ implies that $Z_{n+1} = 0$, $\{\eta_n\}_{n\geq 0}$ is an increasing sequence verifying $\lim_{n\to\infty} \eta_n = \eta$, so by continuity of the function $s \mapsto G_X(s)$, taking the limit in both side of the equation yields (see figure 1),

$$\eta = G_X(\eta)$$

Lemma 2.2. *η is the smallest solution of $\psi = G_X(\psi)$*

Proof. Suppose that $\psi \in [0,1]$ verifies $\psi = G_X(\psi)$. We use induction to prove that for all n, $\eta_n \leq \psi$.

Base case : $\eta_0 = 0 \leq \psi$.

Induction : By induction hypothesis

$$\eta_n \leq \psi$$

The function $s \mapsto G_X(s)$ is increasing on $[0,1]$, therefore

$$G_X(\eta_n) \leq G_X(\psi) = \psi$$

Using equation (4)

$$\eta_{n+1} \leq \psi$$

Proving the induction. As $\lim_{n\to\infty} \eta_n = \eta$, we conclude that

$$\eta \leq \psi$$

□

We continue with the proof of Theorem 2.1.

If $\mathbb{P}[X = 1] = 1$, then for all n, $\mathbb{P}[Z_n = 1] = 1$ so that for all n, $\eta_n = \eta = 0$. If $\mathbb{P}[X \leq 1] = 1$ with $\mathbb{P}[X = 0] = p > 0$, then $\mathbb{P}[Z_n \geq 1] = (1-p)^n$ so that

$$\eta_n = \mathbb{P}[Z_n = 0] = 1 - (1-p)^n \xrightarrow[n\to\infty]{} 1$$

So we can assume that $\mathbb{P}[X \leq 1] < 1$, hence $\mathbb{P}[X > 1] > 0$. Then $G''_X = \mathbb{E}[X(X-1)s^{X-2}] > 0$ and the function $s \mapsto G_X(s)$ is strictly increasing and strictly convex for $s > 0$. Therefore the equation $s = G_X(s)$ has at most two solutions in $[0,1]$. Note that $s = 1$ is a solution for any random variable X, and that $G'_X(1) = \mathbb{E}[X]$. Then there are two cases depending on the value of $\mathbb{E}[X]$ (see figure 2).

- if $G'_X(1) = \mathbb{E}[X] \leq 1$ then there is a unique solution and $\eta = 1$
- if $G'_X(1) = \mathbb{E}[X] > 1$ there are two solutions $s = 1$ and $s = \eta < 1$

□

In many cases we will be interested in the *survival probability*, denoted by $\zeta = 1 - \eta$, the probability that the branching process survives forever,

$$\zeta = \mathbb{P}[Z_n > 0 \ \forall n \geq 0]$$

The total progeny T of the branching process is the total population throughout history, and is defined by

$$T = \sum_{i=0}^{\infty} Z_i \tag{5}$$

We denote by $G_T(s)$ its probability generating function,

$$G_T(s) = \mathbb{E}[e^T]$$

The main result is the following :

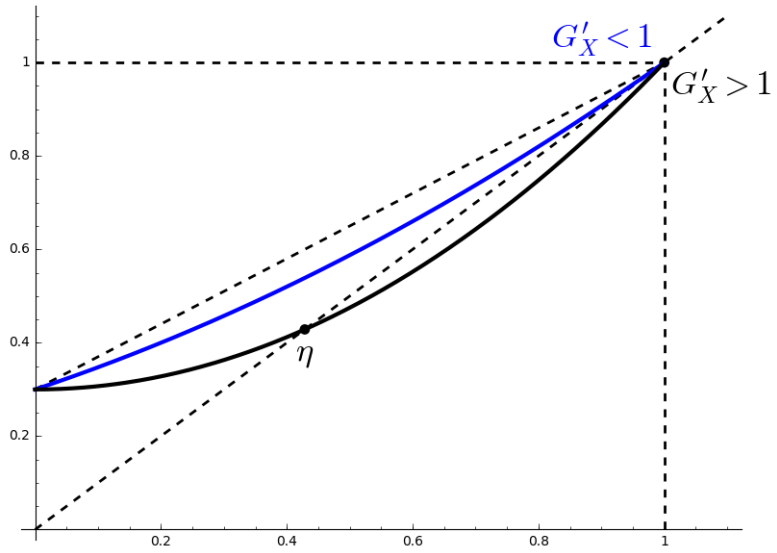


Figure 2: Solutions to $G_X(s) = s$

Theorem 2.3 (Total progeny probability generating function)**.** *For a branching process with i.i.d offspring X having probability generating function $G_X(s) = \mathbb{E}[s^X]$, the probability generating function of the total progeny T satisfies the relation*

$$G_T(s) = sG_X(G_T(s))$$

Proof. We condition on the size of the first generation : For $j = 1, 2, \ldots, Z_1$, the total progeny of the j^{th} child of the initial node satisfies that $\{T_j\}_{j=1}^{Z_1}$ is an i.i.d sequence of random variables. Therefore conditioning on Z_1

$$G_T(s) = \sum_{i=0}^{\infty} p_i \mathbb{E}[s^T \mid Z_1 = i]$$

We use the fact that $T = 1 + \sum_{j=1}^{i} T_j$:

$$G_T(s) = s \sum_{i=0}^{\infty} p_i \mathbb{E}[s^{T_1 + \ldots + T_i}]$$

$$G_T(s) = s \sum_{i=0}^{\infty} p_i G_T(s)^i \text{ the r.v } T_j \text{ are i.i.d}$$

$$G_T(s) = sG_X(G_T(s))$$

□

Theorem 2.4. *For all $n \geq 0$, with $\mu = \mathbb{E}[X]$, the expected number of offspring in the n^{th} generation is given by*

$$\mathbb{E}[Z_n] = \mu^n$$

Proof. Using the fact that

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_{n,i}$$

We condition on Z_{n-1} :

$$\mathbb{E}[Z_n] = \sum_{m=0}^{\infty} \mathbb{P}[Z_{n-1} = m] \cdot \mathbb{E}\left[\sum_{i=1}^{Z_{n-1}} X_{n,i} \mid Z_{n-1} = m\right]$$

The $\{X_{n,i}\}_{n,i\geq 1}$ are a double infinite array of i.i.d random variable. In particular $\{X_{n,i}\}_{i\geq 1}$ is independent from Z_{n-1}. Then

$$\mathbb{E}[Z_n] = \sum_{m=0}^{\infty} \mathbb{P}[Z_{n-1} = m] \cdot m\mathbb{E}\left[X_{n,i}\right] = \mu\mathbb{E}[Z_{n-1}]$$

With $\mathbb{E}[Z_1] = \mathbb{E}[X] = \mu$ we obtain the desired result. □

Applying Markov inequality we directly obtain that

$$\mathbb{P}[Z_n > 0] \leq \mu^n$$

Therefore in the subcritical regime (i.e. when $\mu < 1$), the probability that the population survives up to time n is exponentially small in n.

Corollary 2.5. *For a branching process with i.i.d offspring X having mean offspring $\mu < 1$, the total progeny T verifies*

$$\mathbb{E}[T] = \frac{1}{1-\mu}$$

Proof.

$$\mathbb{E}[T] = \mathbb{E}\left[\sum_{i=0}^{\infty} Z_i\right] = \sum_{i=0}^{\infty} \mu^i = \frac{1}{1-\mu}$$

Or alternatively, conditioning on Z_1:

$$\begin{aligned}
\mathbb{E}[T] &= \sum_{i=0}^{\infty} p_i \mathbb{E}[T \mid Z_1 = i] \\
&= \sum_{i=0}^{\infty} p_i \left(1 + \sum_{j=1}^{i} \mathbb{E}[T_j]\right) \\
&= \sum_{i=0}^{\infty} p_i \left(1 + i\mathbb{E}[T]\right) \\
&= 1 + \mathbb{E}[T] \sum_{i=0}^{\infty} i p_i \\
&= 1 + \mathbb{E}[T]\mu
\end{aligned}$$

Hence

$$\mathbb{E}[T] = \frac{1}{1-\mu}$$

□

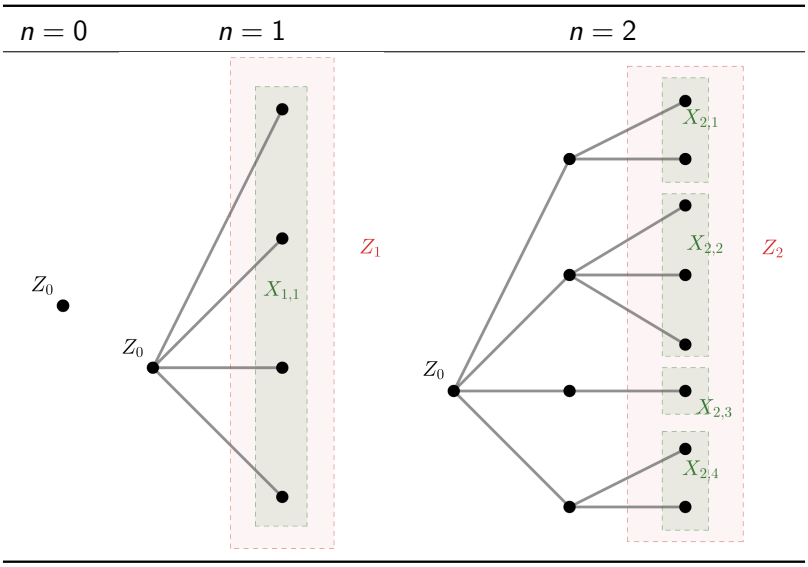


Table 1: Galton-Watson branching process

2.2 Random exploration perspective

The Galton-Watson branching process presented above can be visualised as in table 1.

For all n, i, $X_{n,i} \sim X$ the offspring distribution. For each generation n, its population is given by $Z_n = \sum_{i=0}^{Z_{n-1}} X_{n,i}$, and the total progeny is $T = \sum_{n=0}^{\infty} Z_n$. For future use, we define the random process $(U_i)_{i\in[T]}$ as the number of unexplored individuals after exploring i individuals successively.

This presentation of the branching process is practical to output a distribution of Z_n from X. However for graph purposes, it is also convenient to use a different construction of a branching process.

Let $X_1, X_2, \ldots$ be random variables i.i.d with the same offspring distribution X as $X_{n,i}$. We define by recursion :

$$\begin{cases} S_0 = 1 \\ S_i = S_{i-1} + X_i - 1 = X_1 + X_2 + \ldots + X_i - (i-1) \end{cases} \tag{6}$$

To see how this process compare to our first construction, we look in details to its recursion. We claim that S_t denotes the number of vertices of whom the children have not been explored yet. We call these nodes *active*. At time $t = 0$, we start our exploration process with Z_0 and $S_0 = 1$. Then at each time step i, we select a random active point, explore its children X_i and advance the recursion for S_i. This can be visualised as in table 2.

This process is defined until $S_t = 0$. So let T' be the least t such that $S_t = 0$.

$$T' = \min\{t,\ S_t = 0\}$$

If such a t does not exist, then we set $T' = \infty$.

The above description is equivalent to the definition of the total progeny of our initial branching process. We will demonstrate this fact by proving that the random variable T' is equal in distribution to the total progeny of the branching process T as defined in (5). The proof relies on the formal comparison between S_i's and U_i's:

Lemma 2.6. *The random process $(S_i)_{i\in[T']}$ has the same distribution as the random process $(U_i)_{i\in[T]}$.*

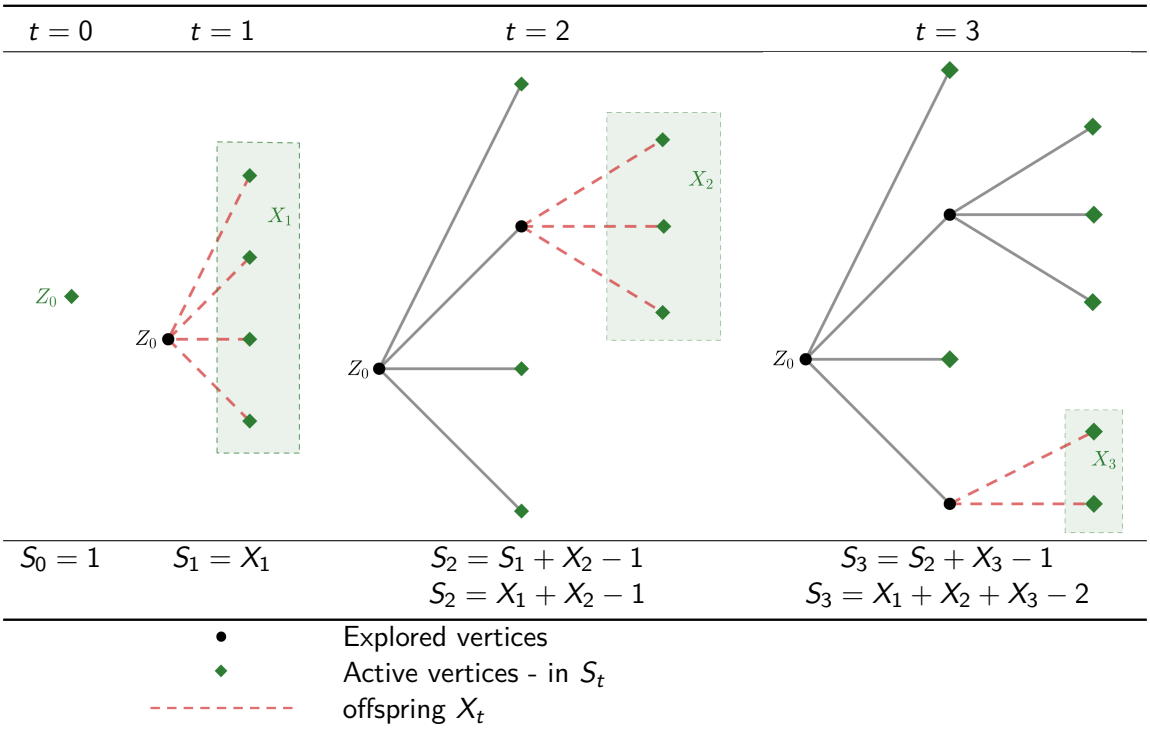


Table 2: Random exploration perspective

Proof. By induction on i : We define our induction hypothesis by

$$\mathcal{H}_i : \{(S_0, \dots, S_i) \text{ has same law as } (U_0, \dots, U_i)\}$$

$S_0 = U_0 = 1$ therefore $\mathcal{H}_0$ is true. Suppose $\mathcal{H}_k$ true for all $k < i$. If $U_{i-1} = 0$, all individuals have been explored, and the total number of explored individuals is equal to the size of the family tree, which is T by definition. Now assume that $U_{i-1} > 0$. We pick a random unexplored individual, and denote it X_i. By independence property of our random process, conditionally on $(U_0, \dots, U_i)$, the distribution of X_i is equal to the distribution of one offspring, e.g. $X_{1,1}$, which is independent from $(U_0, \dots, U_i)$. Then after exploring this i^{th} individual, we've added X_i individual to explore, so that the total number of unexplored individual is

$$U_{i-1} + X_i - 1 = S_{i-1} + X_i - 1$$

Therefore S_i and U_i satisfy the same recursion. $(S_0, \dots, S_i)$ has same law as $(U_0, \dots, U_i)\}$, proving that $\mathcal{H}_i$ is true and concluding our induction. □

We now study the different random trajectories compatible with a total progeny $T = t$, and denote by $H = (X_1, \dots, X_T)$ the *history* of the process up to time T. We include the case where T is infinite, in which case the vector H has infinite length.

$$\mathbb{P}[H = (x_1, \dots, x_t)] = \prod_{i=1}^{t} p_{x_i} \tag{7}$$

Let $(p_k)_{k\geq 0}$ be the offspring distribution of a branching process, conditioned on extinction, and η its

extinction probability. We say that p'_k and p_k are conjugate if they verify

$$p'_k = \eta^{k-1} p_k$$

Recall that $\eta = G_X(\eta)$, so that

$$\sum_i p'_i = \sum_i \eta^{i-1} p_i = \frac{1}{\eta} G_X(\eta) = 1$$

And $(p'_k)_{k\geq 0}$ is a probability distribution.

Theorem 2.7 (Discrete duality principle for branching processes)**.** *Let $(p_k)_{k\geq 0}$ and $(p'_k)_{k\geq 0}$ be a conjugate pair of offspring distributions. The branching process with distribution $(p_k)_{k\geq 0}$ conditioned on extinction, has the same distribution as the branching process with offspring distribution $(p'_k)_{k\geq 0}$.*

Proof. We will show that for every finite history $H = (x_1, \dots, x_t)$, the probability presented in (7) for the branching process $(p_k)_{k\geq 0}$ conditioned to extinction has the same distribution as for the branching process $(p'_k)_{k\geq 0}$.

For any finite t, a finite history implies extinction, therefore

$$\begin{aligned}
\mathbb{P}[H = (x_1, \dots, x_t) \mid \text{ extinction}] &= \frac{\mathbb{P}[\{H = (x_1, \dots, x_t)\} \cap \text{extinction}]}{\mathbb{P}[\text{extinction}]} \\
&= \frac{1}{\eta} \mathbb{P}[H = (x_1, \dots, x_t)] \\
&= \frac{1}{\eta} \prod_{i=1}^{t} p_{x_i}
\end{aligned}$$

Then, by definition of conjugate

$$\begin{aligned}
\mathbb{P}[H = (x_1, \dots, x_t) \mid \text{ extinction}] &= \frac{1}{\eta} \prod_{i=1}^{t} p'_{x_i} \eta^{-(x_i - 1)} \\
&= \frac{1}{\eta} \eta^{t - \sum_{i=1}^{t} x_i} \prod_{i=1}^{t} p'_{x_i}
\end{aligned}$$

Since at extinction $S_t = 0$, then $\sum_{i=1}^{t} x_i = t - 1$ and,

$$\mathbb{P}[H = (x_1, \dots, x_t) \mid \text{ extinction}] = \prod_{i=1}^{t} p'_{x_i} = \mathbb{P}'[H = (x_1, \dots, x_t)]$$

denoting by $\mathbb{P}'$ the distribution of the branching process with offspring distribution $(p'_k)_{k\geq 0}$. □

Theorem 2.10 below will give a specific form of this discrete duality principle in the case of Poisson branching process.

The following theorem shows that for a branching process, if the total progeny is large, then with high probability the process will survive, i.e. it will not die and $T = \infty$.

Theorem 2.8 (Extinction probability with large total progeny)**.** *For a branching process with i.i.d offspring X having mean $\mu > 1$,*

$$\mathbb{P}[k \leq T < \infty] \leq \frac{e^{-\mathcal{I}k}}{1 - e^{-\mathcal{I}}}$$

where the exponential rate $\mathcal{I}$ is given by

$$\mathcal{I} = \sup_{t \leq 0}(t - \log \mathbb{E}[e^{tX}]) > 0$$

Proof. Using the fact that $S_T = 0$ and $S_T = \sum_{i=1}^{T} X_i - (T-1)$ we know that

$$\begin{aligned}
\mathbb{P}[k \leq T < \infty] &\leq \sum_{s=k}^{\infty} \mathbb{P}[S_s = 0] \\
&= \sum_{s=k}^{\infty} \mathbb{P}[X_1 + ... + X_s = s - 1] \\
&\leq \sum_{s=k}^{\infty} \mathbb{P}[X_1 + ... + X_s \leq s]
\end{aligned}$$

For a given s, $\{X_i\}_{i=1}^{s}$ is a sequence of i.i.d random variables with the same probability distribution as X, and having mean $\mu > 1$. Therefore we can apply Chernoff bounds (see second case in A.3) with $a = 1$:

$$\mathbb{P}\left[\sum_{i=1}^{s} X_i \leq sa\right] \leq e^{-s\mathcal{I}},$$

with

$$\mathcal{I} = \sup_{ta \leq 0}(t - \log \mathbb{E}[e^{tX}]) > 0$$

Therefore,

$$P[k \leq T < \infty] \leq \sum_{s=k}^{\infty} e^{-s\mathcal{I}} = \frac{e^{-k\mathcal{I}}}{1 - e^{-\mathcal{I}}}$$

□

We now state a general result on the law of total progeny for a branching process. The proof, omited here, uses the Hitting time theorem and can be found in [Hof16].

Theorem 2.9 (Law of total progeny)**.** *For a branching process with i.i.d offspring X, with $\{X_i\}_{i=1}^{n}$ i.i.d copies of X,*

$$\mathbb{P}[T = n] = \frac{1}{n}\mathbb{P}[X_1 + ... + X_n = n - 1]$$

Notations In the following, we will denote by

- T_{λ}^{Po} the total progeny (or halting time) of a Poisson branching process, with offspring distribution $X \sim Po(\lambda)$

- $T_{n,p}^{Bin}$ the total progeny (or halting time) of a Binomial branching process, with offspring distribution $X \sim Bin(n,p)$
- $T_{n,p}^{Gr}$ the total progeny (or halting time) of a graph branching process, with random graph model $G(n,p)$

2.3 Poisson branching process

In this section, we look at the case where the offspring distribution is Poisson.

For a Poisson random variable X with mean λ, we have the probability generating function of the offspring distribution equal to

$$G_\lambda(s) = \mathbb{E}[s^X] = \sum_{i=0}^{\infty} s^i e^{-\lambda} \frac{\lambda^i}{i!} = e^{\lambda(s-1)}$$

Therefore equation (2) yields

$$\eta = e^{\lambda(\eta-1)} \tag{8}$$

Or equivalently

$$1-\zeta = e^{-\lambda\zeta} \tag{9}$$

For $\lambda \leq 1$, this equation has a unique solution $\eta = 1$ corresponding almost-surely to extinction. If $\lambda > 1$, there are two solutions, the smallest one verifying $\eta \in (0,1)$.

Conditionally on extinction, a Poisson branching process has law $(p'_k)_{k\geq 0}$ given by

$$p'_k = \eta^{k-1} p_k = \eta^{k-1} \frac{e^{-\lambda}\lambda^k}{k!}$$

Using equation (8), we know that $\eta e^{\lambda} = e^{\lambda\eta}$, therefore

$$p'_k = \frac{e^{-(\lambda\eta)}(\lambda\eta)^k}{k!}$$

This distribution is also Poisson, with mean $\mu = \lambda\eta$, and using (8),

$$\mu e^{-\mu} = \lambda\eta e^{-\lambda\eta} = \lambda e^{-\lambda} \tag{10}$$

We will call $\mu < 1 < \lambda$ a conjugate pair if $\mu e^{-\mu} = \lambda e^{-\lambda}$. Note that the function $x \mapsto xe^{-x}$ is increasing then decreasing, with a maximum at $x = 1$. The equation $xe^{-x} = \lambda e^{-\lambda}$ admits two solutions, the trivial one at $x = \lambda > 1$ and $x = \mu < 1$. The discrete duality principle presented in Theorem 2.7 for general branching process takes then a simple form for Poisson branching process.

Theorem 2.10 (Poisson duality principle)**.** *Let $\mu < 1 < \lambda$ be conjugate pairs. The Poisson branching process with mean λ, conditioned on extinction, has the same distribution as a Poisson branching process with mean μ.*

We now investigate further the distribution of the total progeny of a Poisson branching process.

Theorem 2.11 (Total progeny of a Poisson branching process)**.** *For a branching process with i.i.d offspring X, where X has a Poisson distribution with mean λ,*

$$P[T = n] = \frac{(\lambda n)^{n-1}}{n!} e^{-\lambda n}, \quad (n \geq 1) \tag{11}$$

Proof. We use general Theorem 2.9 of total progeny,

$$\mathbb{P}[T = n] = \frac{1}{n} \mathbb{P}[X_1 + ... + X_n = n - 1]$$

We use the fact that for two independent random variables $X \sim Po(\lambda_1)$ and $Y \sim Po(\lambda_2)$, we have $(X + Y) \sim Po(\lambda_1 + \lambda_2)$ Therefore

$$\mathbb{P}[T = n] = \frac{1}{n} \mathbb{P}[Po(n\lambda) = n - 1] = \frac{1}{n} \frac{(n\lambda)^{n-1} e^{-n\lambda}}{(n-1)!}$$

□

Proposition 2.12 (Large but finite Poisson total progeny)**.** *Using Theorem 2.8, we can show that for a branching process with i.i.d offspring X, where X has a Poisson distribution with mean λ,*

$$\mathbb{P}[k \leq T < \infty] \leq e^{-\mathcal{I}_\lambda k},$$

with $\mathcal{I}_\lambda = \lambda - 1 - \log \lambda$.

Proof. Using Theorem 2.8, we have $\mathcal{I} = \sup_{t \leq 0} \left(t - \log \mathbb{E}[e^{tX}] \right)$. The moment generating function of a Poisson random variable $X \sim Po(\lambda)$ is given by

$$\mathbb{E}[e^{tX}] = exp(\lambda(e^t - 1))$$

So that

$$\mathcal{I} = \mathcal{I}_\lambda = \sup_{t \leq 0} \left(t - \lambda(e^t - 1)) \right)$$

The supremum is attained for $t = -\log \lambda$ (by simple derivation) and

$$\begin{aligned} \mathcal{I}_\lambda &= -\log \lambda - \lambda(e^{-\log \lambda} - 1) \\ \mathcal{I}_\lambda &= -\log \lambda - \lambda(\frac{1}{\lambda} - 1) \\ \mathcal{I}_\lambda &= \lambda - 1 - \log \lambda \end{aligned}$$

Therefore

$$\mathbb{P}[k \leq T < \infty] \leq \frac{e^{-\mathcal{I}_\lambda k}}{1 - e^{-\mathcal{I}}} \leq e^{-\mathcal{I}_\lambda k}$$

□

The following theorem gives is a result on the asymptotic of the probability mass function of the total progeny of Poisson branching processes:

Theorem 2.13 (Asymptotic for total progeny for Poisson branching process)**.** *For a branching process with i.i.d offspring X, with $X \sim Po(\lambda)$, as $n \to \infty$,*

$$\mathbb{P}[T = n] \sim \frac{1}{\lambda\sqrt{2\pi n^3}} e^{-\mathcal{I}_\lambda n} \tag{12}$$

Proof. We use Stirling's approximation:

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1 + O(n^{-1}))$$

From the total progeny of a Poisson branching process (equation 11):

$$\begin{aligned}
\mathbb{P}[T = n] &= \frac{(\lambda n)^{n-1}}{n!} e^{-\lambda n} \\
&\sim \frac{1}{\lambda n} \frac{1}{\sqrt{2\pi n}} \lambda^n e^n e^{-\lambda n} \\
&\sim \frac{1}{\lambda\sqrt{2\pi n^3}} e^{-(\lambda - 1 - \log\lambda)n} \\
&\sim \frac{1}{\lambda\sqrt{2\pi n^3}} e^{-\mathcal{I}_\lambda n}
\end{aligned}$$

□

Note that for any $\lambda \neq 1$, $\mathcal{I}_\lambda$ is strictly positive, hence $\mathbb{P}[T = n]$ tends to 0 exponentially. This gives us a bound on the tail of the distribution:

$$\mathbb{P}[T \geq k] < e^{-k(\mathcal{I}_\lambda + o(1))} \tag{13}$$

In particular, for $\lambda = 1 - \varepsilon$, for $\varepsilon \to 0^+$, and $A \to \infty$:

$$\mathbb{P}[T_{1-\varepsilon} > A\varepsilon^{-2}] < \varepsilon e^{-(1+o(1))A/2} \tag{14}$$

2.4 The binomial branching process

In a random graph with probability of having an edge $= \lambda/n$, the total number of vertices incident to a particular vertex has a binomial distribution $Bin(n, \lambda/n)$. Using coupling, we can relate the total progeny of a Poisson branching process to that of a binomial branching process with parameters n and success probability λ/n.

Theorem 2.14 (Poisson limit for binomial random variables)**.** *Let $\{I_i\}_{i=1}^n$ be independent with $I_i \sim Be(p_i)$, and let $\lambda = \sum_{i=1}^n p_i$. Let $X = \sum_{i=1}^n I_i$ and $Y \sim Po(\lambda)$. Then there exists a coupling $(\widehat{X}, \widehat{Y})$ of (X, Y) such that*

$$\mathbb{P}[\widehat{X} \neq \widehat{Y}] \leq \sum_{i=1}^n p_i^2 \tag{15}$$

Consequently, for every $\lambda \geq 0$ and $n \in \mathbb{N}$, there exists a coupling $(\widehat{X}, \widehat{Y})$ with $\widehat{X} \sim Bin(n, \lambda/n)$ and $\widehat{Y} \sim Po(\lambda)$ such that

$$\mathbb{P}[\widehat{X} \neq \widehat{Y}] \leq \frac{\lambda^2}{n} \tag{16}$$

Proof. Let $J_i \sim Po(p_i)$ all independent. Let $(\widehat{I}_i, \widehat{J}_i)$ be the maximal coupling of (I_i, J_i)*, satisfying for a given x

$$\mathbb{P}[\widehat{I}_i = \widehat{J}_i = x] = \min\{\mathbb{P}[I_i = x], \mathbb{P}[J_i = x]\} = \begin{cases} 1 - p_i & \text{for } x = 0 \\ p_i e^{-p_i} & \text{for } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Therefore,

$$\mathbb{P}[\widehat{I}_i = \widehat{J}_i] = \sum_x \min\{\mathbb{P}[I_i = x], \mathbb{P}[J_i = x]\}$$

Using $1 - p_i \leq e^{-p_i}$,

$$\mathbb{P}[\widehat{I}_i \neq \widehat{J}_i] = 1 - (1 - p_i) - p_i e^{-p_i} = p_i(1 - e^{-pi}) \leq p_i^2$$

We conclude by definition $\widehat{X} = \sum_{i=1}^n \widehat{I}_i$ and $\widehat{Y} = \sum_{i=1}^n \widehat{J}_i$, so that $\widehat{X} \sim X$ and $\widehat{Y} \sim Y$ (using independence of J_i r.v). Using the union bound,

$$\mathbb{P}[\widehat{X} \neq \widehat{Y}] \leq \mathbb{P}\left[\bigcup_{i=1}^n \{\widehat{I}_i \neq \widehat{J}_i\}\right] \leq \sum_{i=1}^n \mathbb{P}[\widehat{I}_i \neq \widehat{J}_i] \leq \sum_{i=1}^n p_i^2$$

□

We can now relate the total progeny of a Poisson branching process to that of a binomial branching process.

Theorem 2.15 (Poisson approximation for BR)**.**

$$\mathbb{P}[T_{n,p}^{Bin} \geq k] = \mathbb{P}[T_\lambda^{Po} \geq k] + e_k(n)$$

with

$$|e_k(n)| \leq \frac{2\lambda^2}{n} \sum_{s=1}^{k-1} \mathbb{P}_\lambda[T_\lambda^{Po} \geq k]$$

In particular,

$$|e_k(n)| \leq \frac{2k\lambda^2}{n}$$

Proof. Let X_i^{Bin} and X_i^{Po} be the random variable determining their respective branching process, as per the recursion defined in (6). For each i, we use the coupling presented in Theorem 2.14 with $X_i^{Bin} \sim Bin(n, \lambda/n)$, $X_i^{Po} \sim Po(\lambda)$ and

$$\mathbb{P}\left[X_i^{Bin} \neq X_i^{Po}\right] \leq \frac{\lambda^2}{n} \tag{17}$$

Note that (using the joint probability),

$$\begin{aligned}\left|\mathbb{P}[T^{Bin} \geq k] - \mathbb{P}[T^{Po} \geq k]\right| = \Big|&\mathbb{P}[T^{Bin} \geq k, T^{Po} \geq k] + \mathbb{P}[T^{Bin} \geq k, T^{Po} < k] \\ &- \mathbb{P}[T^{Po} \geq k, T^{Bin} \geq k] - \mathbb{P}[T^{Po} \geq k, T^{Bin} < k]\Big|\end{aligned}$$

*see [LP17] chapter 5 for a full introduction to coupling

So that

$$\left|\mathbb{P}[T^{Bin} \geq k] - \mathbb{P}[T^{Po} \geq k]\right| \leq \max\left\{\mathbb{P}[T^{Bin} \geq k,\, T^{Po} < k], \mathbb{P}[T^{Po} \geq k,\, T^{Bin} < k]\right\}$$

If $T^{Bin} \geq k$ while $T^{Po} < k$, then necessarily there exists some $s < k$ such that $X_s^{Bin} \neq X_s^{Po}$. Using the union bound, we can split the probability depending on the first s where the branching processes disagree :

$$\mathbb{P}[T^{Bin} \geq k,\, T^{Po} < k] \leq \sum_{s=1}^{k-1} \mathbb{P}[X_s^{Bin} \neq X_s^{Po},\ \ T^{Bin} \geq k,\ \forall i < s\ X_i^{Bin} = X_i^{Po}]$$

For a given s, if for all $i < s$, $X_i^{Bin} = X_i^{Po}$ while $T^{Bin} \geq k > s$ then this implies that $T^{Po} \geq s$.

$$\mathbb{P}[T^{Bin} \geq k,\, T^{Po} < k] \leq \sum_{s=1}^{k-1} \mathbb{P}[T^{Po} \geq s,\ X_s^{Bin} \neq X_s^{Po}]$$

Now the event $\{T^{Po} \geq s\}$ depends only on $X_1^{Po}, \dots, X_{s-1}^{Po}$ and is therefore independent from $X_s^{Bin} = X_s^{Po}$. Then

$$\mathbb{P}[T^{Bin} \geq k,\, T^{Po} < k] \leq \sum_{s=1}^{k-1} \mathbb{P}[T^{Po} \geq s]\mathbb{P}[X_s^{Bin} \neq X_s^{Po}]$$

Using (17),

$$\mathbb{P}[T^{Bin} \geq k,\, T^{Po} < k] \leq \frac{\lambda^2}{n} \sum_{s=1}^{k-1} \mathbb{P}[T^{Po} \geq s]$$

By symmetry we also obtain that

$$\mathbb{P}[T^{Po} \geq k,\, T^{Bin} < k] \leq \frac{\lambda^2}{n} \sum_{s=1}^{k-1} \mathbb{P}[T^{Po} \geq s]$$

and

$$\left|\mathbb{P}[T^{Bin} \geq k] - \mathbb{P}[T^{Po} \geq k]\right| \leq \frac{\lambda^2}{n} \sum_{s=1}^{k-1} \mathbb{P}[T^{Po} \geq s]$$

□

3. Erdős-Rényi Model

Before presenting the Erdős-Rényi model, we start by giving some graph theory definitions used in this document.

Definition 3.1 (order and size)**.** Let $G = (V, E)$ be a graph with vertices set S and edges set E. The *order* of G is the number of vertices in G, usually denoted $n = |V|$. The *size* of G is the number of edges in G, usually denoted $m = |E|$.

Definition 3.2 (simple graph and multigraph)**.** Let $G = (V, E)$ be a graph. We say that the graph is *simple* if there is neither loop (edge starting and ending at the same vertex), nor multi-edges (several edges with the same two endpoints). Otherwise the graph is said to be a *multigraph*.

Definition 3.3 (Connected component)**.** Given a graph G, the connected component containing a vertex v, denoted $\mathcal{C}(v)$, is the set of all vertices of G that can be reached from v using existing edges. The largest connected component is equal to any connected component $\mathcal{C}(v)$ such that $|\mathcal{C}(v)|$ is maximal, so that:

$$|\mathcal{C}_{max}| = \max\{|\mathcal{C}(v)|,\ v \in [n]\}$$

Definition 3.4 (Complexity)**.** Given a graph G, the complexity of one connected component $\mathcal{C}$ is its number of edges m minus its number of vertices n :

$$c(\mathcal{C}) = m - n$$

If $c = -1$ or $c = 0$, the component is said to be *simple*. Otherwise it is *complex*.

Remark 3.5. For connected graphs, $c(G) = -1$ implies that G is a tree, while $c(G) = 0$ implies that G has exactly one cycle.

In the following, we will denote L_k as the k^{th} largest connected component, so that $L_1 = |\mathcal{C}_{max}|$.

3.1 Presentation and properties

In this section, we introduce the Erdős-Rényi model $G(n, p)$ and state some first properties of random graphs in this model.

A random graph from $G(n, p)$ has vertex set $[n] = \{1, 2, \dots, n\}$ and each edge is present with probability p (called the edge probability), independently of all other edges.

In other words, if G is a random graph with distribution $G(n, p)$, and H is a specific graph with n vertices and m edges then,

$$\mathbb{P}[G = H] = p^m(1-p)^{\binom{n}{2}-m}$$

For a graph $G \in G(n, p)$, the total number of edges is the sum of $\binom{n}{2}$ i.i.d variables $X \sim Be(p)$, therefore it follows a distribution $Bin(\binom{n}{2}, p)$. In particular :

$$\begin{cases} \mathbb{E}[\text{number of edges}] &= \binom{n}{2}p \\ \mathbb{V}ar[\text{number of edges}] &= \binom{n}{2}p(1-p) \end{cases}$$

We also have the following (well-known) lemma (see for instance [Chv91]), giving us, for each i, the number n_i of vertices of degree i,

Proposition 3.6. *For a random graph $G \in G(n, p)$ with $p = c/n$, for $i \leq O(\log n / \log\log n)$, a.s.*

$$n_i = \frac{c^i}{i!}e^{-c}n + o(n),$$

and $n_i = 0$ otherwise.

This is easily shown using the fact that for any vertex, its degree follows a binomial distribution $Bin(n-1, c/n)$, which we can approximate with a Poisson random variable $Po(c)$.

3.2 Exploration - Graph branching process

In this section we define how to determine the connected component containing a given vertex v, in relation to the random walks perspective of a branching process.

Let $G \in G(n,p)$ be a random graph on n vertices. Seeing our graph as a branching process, we recall random exploration perspective : during the exploration process, vertices can have three statuses: active, unexplored or explored.

At each time t, we denote by S_t the number of active vertices. At time $t = 0$, v is active, all other vertices are unexplored, so that $S_0 = 1$. The exploration process works as follow, for each time step t:

(1) select w_t at random among active vertices (e.g. assign an ordering to all vertices $\{1, \dots, n\}$ and select the smallest one)

(2) Explore all unexplored vertices w' such that $w_t w'$ is an edge of G. Set all w' to active.

(3) Set w_t as explored.

Let T be the least t such that $S_t = 0$.

$$T = \min\{t,\ S_t = 0\}$$

At time T, the exploration process terminates, and $\mathcal{C}(v)$ is the set of all explored vertices, so that $|\mathcal{C}(v)| = T$.

Let X_t be the number of vertices turning active (from unexplored) due to the exploration of w_t, i.e. $X_t = |\{w',\ w_t w' \in E(G)\}|$. At each time t, the number of active vertices increase by X_t and decrease by 1 (w_t is set as explored). Hence we have the recursion

$$\begin{cases} S_0 = 1 \\ S_i = S_{i-1} + X_i - 1 = X_1 + X_2 + \dots + X_i - (i-1) \end{cases}$$

In $G(n,p)$, each edge is independent from all other edges. Therefore at each time t, each unexplored vertex w' has a probability p to be linked to w_t, hence has a probability p to become active. Noting that after t steps they are t explored vertices and S_t active vertices, then conditionally on S_t:

$$X_{t+1} \sim Bin(n - t - S_t, p) \tag{18}$$

Let N_t the number of unexplored vertices at time t, we have $N_t = n - t - S_t$ and $N_t = N_{t-1} - X_t$ so that $N_t \sim Bin(N_{t-1}, 1-p)$ and we can show by induction that

$$N_t \sim Bin(n-1, (1-p)^t) \qquad \text{for } 0 \le t \le n$$

If $T = t$ it is necessary that $N_t = n - t$, hence

$$\mathbb{P}[|\mathcal{C}(v)| = t] \le \mathbb{P}\left[Bin(n-1,(1-p)^t) = n - t\right] \tag{19}$$

Note that $X \sim Bin(m,q)$ if and only $Y = m - X \sim Bin(m, 1-q)$. Therefore as $N_t = n - t - S_t \sim Bin(n-1,(1-p)^t)$ then

$$S_t + (t-1) \sim Bin(n-1, 1-(1-p)^t)$$

And equation 19 is equivalent to

$$\mathbb{P}[|\mathcal{C}(v)| = t] \le \mathbb{P}\left[Bin(n-1,1-(1-p)^t) = t - 1\right] \tag{20}$$

3.3 Relations between processes

Let $p = c/n$. Given equation (18), we know that $X_1 \sim Bin(n-1, c/n)$. Then as $n \to \infty$, X_1 approaches the distribution of a Poisson with parameter c . Moreover, the same holds for X_t as long as $N_{t-1} \sim o(n)$, i.e. *the exploration of $\mathcal{C}(v)$ mimics the Poisson branching process with mean c as long as the number of vertices found is not too large*. This will allow for a precise description of $\mathcal{C}(v)$ is the subcritical regime $c < 1$. But in the supercritical regime $c > 1$, this relation between the two processes must fail : There must be an *ecological limitation*. In the supercritical regime, the Poisson branching process is infinite with positive probability, while any component is tautologically bounded in size by n. In the graph process, as the number N_t of unexplored vertices decreases, so does $\mathbb{E}[X_t]$ the number of vertices added to the active list. Eventually the drift of S_t lowers to negative values, causing the process to halt.

We will now investigate the relations between connected components and binomial branching processes.

3.3.1 Stochastic domination of connected components

We start by proving that each connected component is bounded from above by the total progeny of a branching process with binomial offspring distribution.

Theorem 3.7. *Let $T_{n,p}^{Bin}$ be the total progeny of a binomial branching process $Bin(n,p)$, and v the root node of the graph branching process. For each $k \geq 1$,*

$$\mathbb{P}[|\mathcal{C}(v)| \geq k] = \mathbb{P}[T_{n,p}^{Gr} \geq k] \leq \mathbb{P}[T_{n,p}^{Bin} \geq k]$$

Proof. Let $N_i = n - i - S_i$ be the number of unexplored vertices after i explorations, so that, given equation (18), conditionally on N_{i-1},

$$X_i \sim Bin(N_{i-1}, p)$$

Let $Y_i \sim Bin(n - N_{i-1}, p)$ and $A_i = X_i + Y_i$. Then, conditionally on all $\{X_j\}_{j=1}^{i-1}$, the $(A_j)_{j\geq 1}$ are i.i.d. and

$$A_i \sim Bin(n, p)$$

Now let $B_i = A_1 + ... + A_i - (i-1)$. We have

$$\begin{cases} S_i = X_1 + ... + X_i - (i-1) \\ B_i = A_1 + ... + A_i - (i-1) \end{cases} \quad \text{and } \forall i, A_i \geq X_i \text{ since } Y_i \geq 0$$

Therefore

$$\mathbb{P}[|\mathcal{C}(v)| \geq k] = \mathbb{P}[S_t > 0,\ \forall t \leq k-1] \leq \mathbb{P}[B_t > 0,\ \forall t \leq k-1] = \mathbb{P}[T_{n,p}^{Bin} \geq k]$$

□

Remark 3.8. Using corollary 2.5, this theorem implies that for $\mu = (n-1)p < 1$,

$$\mathbb{E}[|\mathcal{C}(v_0)|] \leq \frac{1}{1-\mu}$$

3.3.2 Lower bound for connected component

We now give a lower bound for the size of connected components, using the total progeny of a branching process with binomial offspring distribution, now with parameters $Bin(n-k,p)$.

Theorem 3.9. *For each $k \in [n]$, let $T^{Bin}_{n-k,p}$ be the total progeny of a binomial branching process $Bin(n-k,p)$, and v the root node of the graph branching process. Then*

$$\mathbb{P}[|\mathcal{C}(v)| \geq k] = \mathbb{P}[T^{Gr}_{n,p} \geq k] \geq \mathbb{P}[T^{Bin}_{n-k,p} \geq k]$$

Note that as the binomial distribution used in this lower bound depends on k, we do not have here a stochastic lower bound on $|\mathcal{C}(v)|$.

Proof. Recall that in our graph branching process, N_i denotes the number of unexplored vertices after i explorations. Denote the time τ_k by

$$\tau_k = \min\{t,\ N_t < n-k\}$$

At each exploration step, at least one vertex is moved from being unexplored to being active. Therefore $\tau_k < k$. Let $\widehat{X}_i$ be a sequence of i.i.d random variables $Bin(n-k,p)$, and conditionally on N_{i-1}, for $i \leq \tau_k + 1$, let

$$X_i = \widehat{X}_i + Y_i$$

So that $Y_i \sim Bin(N_{i-1} - (n-k), p)$. Recall that we want to bound

$$\mathbb{P}[|\mathcal{C}(v)| \geq k] = \mathbb{P}[S_t > 0,\ \forall t \leq k-1]$$

Note that $\tau_k < k$ implies that

$$\{S_t > 0,\ \forall t \leq k-1\} \subseteq \{S_t > 0,\ \forall t \leq \tau_k\} \tag{21}$$

Note also that $S_{\tau_k} = n - \tau_k - N_{\tau_k}$. By definition of τ_k, $N_{\tau_k} \leq n-k$ so that

$$S_{\tau_k} \geq k - \tau_k$$

By definition of our graph branching process (recall that S_t represent the number of active vertices) S_t decreases at most by 1 at each exploration step, so that we need at least $k - \tau_k$ additional steps for S_t to be null:

$$\begin{cases} S_{\tau_k} \geq k - \tau_k \\ \forall t \leq \tau_k,\ S_t > 0 \end{cases} \Rightarrow \forall t \leq k-1,\ S_t > 0$$

And

$$\{S_t > 0,\ \forall t \leq k-1\} \supseteq \{S_t > 0,\ \forall t \leq \tau_k\}$$

Combining with equation (21),

$$\mathbb{P}[|\mathcal{C}(v)| \geq k] = \mathbb{P}\left[S_t > 0,\ \forall t \leq k-1\right] = \mathbb{P}\left[S_t > 0,\ \forall t \leq \tau_k\right] \tag{22}$$

Now let

$$\widehat{S}_i = \widehat{X}_1 + \ldots + \widehat{X}_i - (i-1)$$

Because $Y_i \geq 0$ a.s., then $\widehat{X}_i \leq X_i$ a.s. and $\widehat{S}_i \leq S_i$ a.s. And finally

$$\mathbb{P}\left[S_t > 0,\ \forall t \leq \tau_k\right] \geq \mathbb{P}\left[\widehat{S}_t > 0,\ \forall t \leq \tau_k\right] \geq \mathbb{P}\left[\widehat{S}_t > 0,\ \forall t \leq k-1\right] = P[T^{Bin}_{n-k,p} \geq k]$$

□

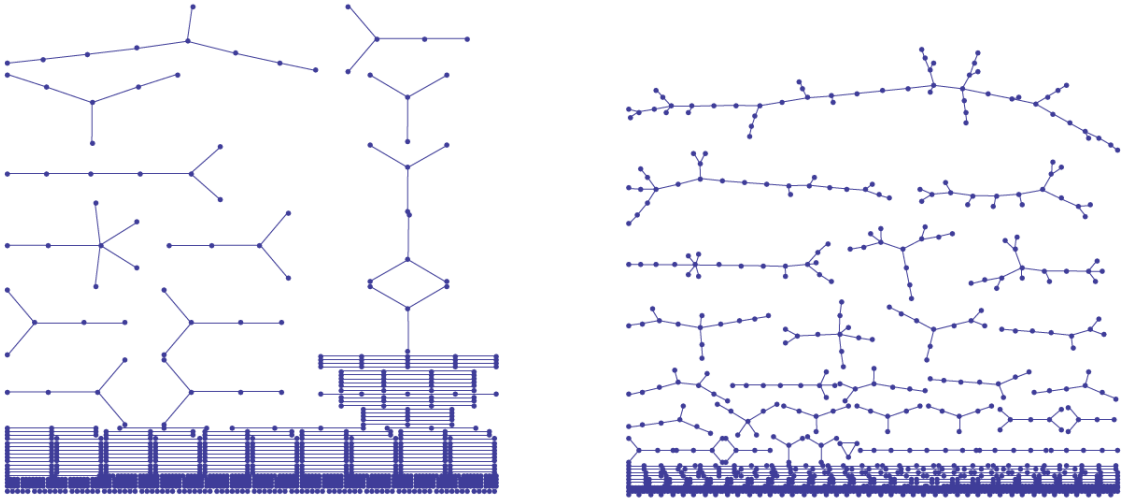


Figure 3: Realizations of Erdős-Rényi random graphs with 1000 elements and edge probabilities $c/1000$ with $c = 0.5$ and $c = 0.9$ respectively. Source [Hof16]

3.4 Overview of results

Before entering into detailed results, we give here an overview of the study of the largest components in $G(n, p)$. This usually includes five regimes:

Very subcritical We use a coarse parametrization : $p = c/n$ with $c < 1$. In this regime (see figure 3),

- all components are simple,
- $L_1 = \Theta(\ln n)$,
- $L_k \sim L_1$ for all k.

Barely subcritical We use a fine parametrization $p = (1 - \varepsilon)/n$ with $\varepsilon = \lambda n^{-1/3}$, $\varepsilon = o(1)$ and $\lambda \to \infty$. In this regime,

- all components are simple,
- $L_1 = \Theta(n^{2/3}\lambda^{-2} \ln \lambda)$,
- $L_k \sim L_1$ for all k.

Critical window We use a fine parametrization with λ real **constant** : $p = 1/n + \lambda n^{-4/3}$. In this regime(see figure 4)„

- for a given λ, the sizes of the largest components are of the form $cn^{2/3}$ with a distribution over the constant c with support in $\mathbb{R}^+$. Hence, for any λ and any $c > 0$, asymptotically, the probability that the largest component is bigger (or smaller) than $cn^{2/3}$ is strictly positive.
- When λ is largely negative, then with high probability most (but not all) of the components are trees, and many components will have about the same size.
- When λ grows, the dominant component starts to emerge, of non-simple complexity, and the second largest component is simple and far smaller.

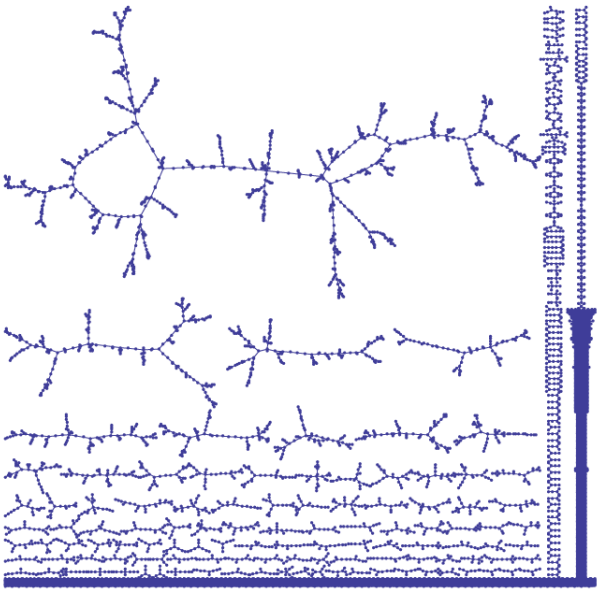


Figure 4: Realizations of Erdős-Rényi random graphs with 10000 elements and edge probabilities $c/10000$ with $c = 1$. Source [Hof16]

Barely supercritical We use a fine parametrization $p = (1+\varepsilon)/n$ with $\varepsilon = \lambda n^{-1/3}$, $\varepsilon = o(1)$ and $\lambda \to \infty$. In this regime,

- $L_1 \sim 2\lambda n^{2/3}$,
- the largest component has complexity approaching infinity, all other are simple,
- $L_2 = \Theta(n^{2/3}\lambda^{-2} \ln \lambda)$.

Very supercritical We use a coarse parametrization : $p = c/n$ with $c > 1$. In this regime(see figure 5),

- $L_1 \sim \zeta n$ with ζ the smallest solution to $e^{-c\zeta} = 1 - \zeta$. ζ is the survival probability of our graph branching process (see equation (9)).
- The largest component has complexity approaching infinity, all other are simple,
- $L_2 = \Theta(\ln n)$.

Note that in the barely supercritical regimes (equiv. very supercritical) the size of the second largest component is equivalent to the size of the largest one in barely subcritical (equiv. very subcritical),

$$L_2^{super} \sim L_1^{sub}$$

This will give birth to the discrete duality principal of the Erdős-Rényi model.

Explaining the parametrization

The coarse parametrization separates the studies into three intuitive regimes. A subcritical $p = c/n$ with $c < 1$, a supercritical $p = c/n$ with $c > 1$, and a critical one $p = 1/n$. However the fine parametrization relies on a non-intuitive factor of $4/3$, defining 5 regimes (see table 3).

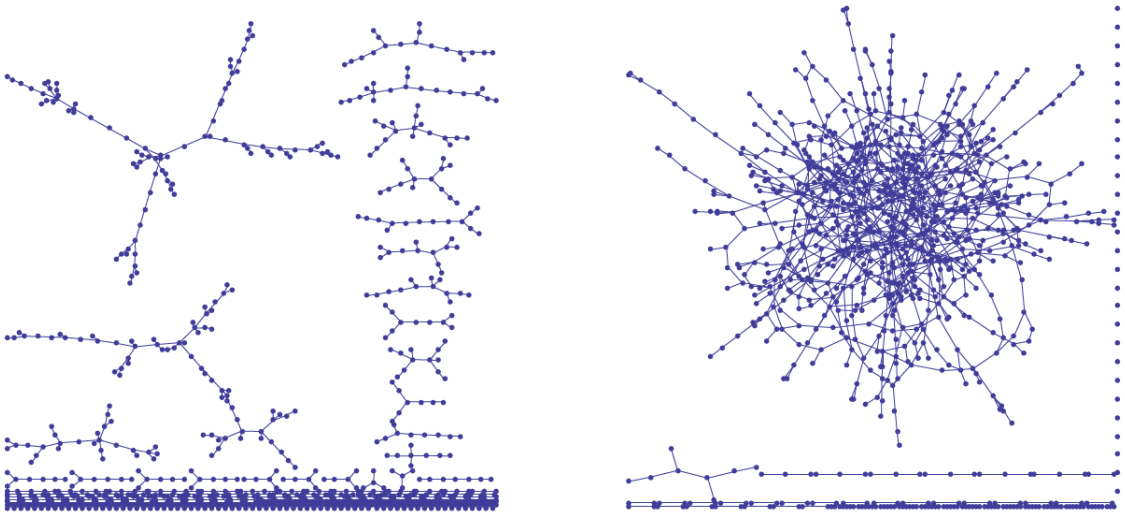


Figure 5: Realizations of Erdős-Rényi random graphs with 1000 elements and edge probabilities $c/1000$ with $c = 1.1$ and $c = 2$ respectively. Source [Hof16]

Regime	Very subcritical	Barely subcritical	Critical window	Barely supercritical	Very supercritical
p	c/n	$1/n - \lambda n^{-4/3}$	$1/n \pm \lambda n^{-4/3}$	$1/n + \lambda n^{-4/3}$	c/n
	$c < 1$ constant	$\lambda \to \infty$	λ constant	$\lambda \to \infty$	$c > 1$ constant

Table 3: the fine parametrization

In their book [AS92], Alon and Spencer gave a quick heuristic explanation for this factor: Let $p = (1+\varepsilon)/n$ with $\varepsilon \to 0$, and consider the Poisson branching process $T = T^{Po}_{1+\varepsilon}$. It is infinite with probability $\sim 2\varepsilon$ and otherwise the probability that T exceeds $A\varepsilon^{-2}$ decreases exponentially in A.

We will see in the barely supercritical study (see section 3.6.2) that the case where T is infinite translates into our graph with a dominant component of size $\sim 2\varepsilon n$. In order for this to hold, we need a distinction between the small components of size up to $\sim \varepsilon^{-2}$ and the dominant one, hence $2\varepsilon n \gg \varepsilon^{-2}$. This heuristically leads to a breakpoint at $\varepsilon = n^{-1/3}$. If $\varepsilon \gg n^{-1/3}$, we have the distinction between small and dominant components, and we are in the supercritical regime.

3.5 Subcritical regime

In the subcritical regime, we will bound by above the size of the largest connected component of $G(n,p)$.

3.5.1 Very subcritical phase

Let $p = c/n$ with $c < 1$. Recall that by theorem 3.7, each connected component is bounded from above by the total progeny of a branching process with binomial offspring distribution:

$$\mathbb{P}[|\mathcal{C}(v)| \geq k] \leq \mathbb{P}[T^{Bin}_{n,p} \geq k]$$

Using the Poisson approximation (Theorem 2.15),

$$\mathbb{P}[|\mathcal{C}(v)| \geq k] \leq (1 + o(1))\mathbb{P}[T^{Po}_{c} \geq k]$$

In the section devoted to the Poisson branching process we have seen that this values decreases exponentially in k (see equation (13)):

$$\mathbb{P}[T_c^{Po} \geq k] < e^{-k(\mathcal{I}_c + o(1))}$$

With $\mathcal{I}_c = c - 1 - \log c > 0$

Then for any $\varepsilon > 0$, there is exist a large enough constant K, with $k = K \ln n$, such that

$$\mathbb{P}[|\mathcal{C}(v)| \geq K \ln n] \leq n^{-\mathcal{I}_c K} < n^{-(1+\varepsilon)}$$

This holds for any of the n possible choices of initial vertex v for the exploration process, therefore we can upper bound the probability to have any component of size greater then $K \ln n$:

$$\mathbb{P}[\exists v,\ |\mathcal{C}(v)| \geq K \ln n] < n n^{-(1+\varepsilon)} = n^{-\varepsilon} \to 0$$

Therefore

$$\boxed{L_1 = O(\ln n) \quad a.a.s}$$

3.5.2 Barely subcritical regime

Let $p = (1-\varepsilon)/n$ with $\varepsilon = \lambda n^{-1/3}$, $\varepsilon = o(1)$ and $\lambda \to \infty$.

Let $\chi_v(k)$ be the characteristic random variable for $\mathcal{C}(v)$ having at least k vertices,

$$\chi_v(k) = \begin{cases} 1 & \text{If } |\mathcal{C}(v)| \geq k \\ 0 & \text{If } |\mathcal{C}(v)| < k \end{cases}$$

As seen above, Theorem 3.7 and the Poisson approximation give,

$$\mathbb{P}[|\mathcal{C}(v)| \geq k] \leq (1 + o(1))\mathbb{P}[T_{1-\varepsilon}^{Po} \geq k]$$

We recall equation (14) on the tail distribution,

$$\mathbb{P}[T_{1-\varepsilon} > A\varepsilon^{-2}] < \varepsilon e^{-(1+o(1))A/2}$$

Let $k = K\varepsilon^{-2} \ln \lambda = K n^{2/3} \lambda^{-2} \ln \lambda$, then

$$\mathbb{P}[T_{1-\varepsilon} > k] < \varepsilon e^{-(1+o(1))K \ln \lambda / 2} = \varepsilon \lambda^{-(1+o(1))K/2}$$

For any $\theta > 0$, for sufficiently large constant K,

$$\mathbb{P}[T_{1-\varepsilon}^{Poi} \geq k] \leq \varepsilon \lambda^{-3(1+\theta)}$$

Let $Z = \sum_v \chi_v(k)$ be the number of vertices v in components of size at least k, and Y the number of components of size at least k. Then

$$\mathbb{E}[Z] = n\mathbb{E}[\chi_v(k)] \leq n\varepsilon\lambda^{-3(1+\theta)} = n^{2/3}\lambda^{-2(1+\theta)}$$

As $kY \leq Z$ (as each Y component has size at least k), using the fact that $\lambda \to \infty$,

$$\mathbb{E}[Y] \leq k^{-1}\mathbb{E}[Z] \leq (K \ln \lambda)^{-1} \lambda^{-2\theta} \to 0$$

Therefore a.a.s $Y = 0$ and so

$$\boxed{L_1 \leq K n^{2/3} \lambda^{-2} \ln \lambda \quad \text{a.a.s}}$$

3.6 Supercritical regime

In the supercritical regime, there are two main results on the largest component, first its existence and then its uniqueness.

3.6.1 Very supercritical phase

Let $p = c/n$ with $c > 1$ constant,

- Let $\zeta = \zeta(c)$ be the smallest positive real solution to $e^{-c\zeta} = 1 - \zeta$. ζ is the survival probability of a Poisson branching process with parameter c (see equation (9)).
- Let δ be an arbitrarily small constant and K an appropriately large constant
- Set $k_n = K \ln n$, $L^- = (\zeta - \delta)n$ and $L^+ = (\zeta + \delta)n$

We call a connected component $\mathcal{C}(v)$

- **small** if $|\mathcal{C}(v)| < k_n$
- **giant** if $L^- < |\mathcal{C}(v)| < L^+$
- **awkward** otherwise

No Middle Ground

Theorem 3.10. *Let $\alpha < \zeta$, $\alpha = \zeta - \delta$ such that $L^- = \alpha n$. For sufficiently large K, with probability at least $(1 - n^{-\theta})$, for some arbitrarily large θ detailed below, there is no connected component of size in between k_n and αn.*

Note: In the following, the sums over all possible size should run from $\lfloor k_n \rfloor$ to $\lfloor \alpha n \rfloor$. In order to avoid heavy notation, we will abuse the notations writing k_n and αn. This does not affect the results.

Proof. Recall that, from equation (20), for any fixed vertex v,

$$\mathbb{P}[|\mathcal{C}(v)| = t] \leq \mathbb{P}\left[Bin\left(n-1, 1-\left(1-\frac{c}{n}\right)^t\right) = t-1\right]$$

Therefore

$$\begin{aligned}
\mathbb{P}[k_n \leq |\mathcal{C}(v)| \leq \alpha n] &\leq \sum_{t=k_n}^{\alpha n} \mathbb{P}\left[Bin\left(n-1, 1-\left(1-\frac{c}{n}\right)^t\right) = t-1\right] \\
&\leq \sum_{t=k_n}^{\alpha n} \mathbb{P}\left[Bin\left(n-1, 1-\left(1-\frac{c}{n}\right)^t\right) \leq t-1\right] \\
&\leq \sum_{t=k_n}^{\alpha n} \mathbb{P}\left[Bin\left(n, 1-\left(1-\frac{c}{n}\right)^t\right) \leq t\right]
\end{aligned}$$

Using the fact that $1 - x \leq e^{-x}$:

$$\mathbb{P}[k_n \leq |\mathcal{C}(v)| \leq \alpha n] \leq \sum_{t=k_n}^{\alpha n} \mathbb{P}\left[Bin\left(n, 1 - \left(1 - e^{-ct/n}\right)\right) \leq t\right]$$

Note that for $x = \zeta$, $1 - e^{-c\zeta} = \zeta$, hence for $t < \alpha n < \zeta n$, we have $t/n < 1 - e^{-ct/n}$, and we can apply Chernoff bound for $X \sim \mathbb{E}[Ber(1 - e^{-ct/n})]$:

$$\mathbb{P}[k_n \leq |\mathcal{C}(v)| \leq \alpha n] \leq \sum_{t=k_n}^{\alpha n} e^{-n\mathcal{I}(t/n)}$$

With

$$\mathcal{I}(t/n) = \sup_{s \leq 0}\left\{s\frac{t}{n} - \log \mathbb{E}[e^{sX}]\right\}$$

So that

$$\begin{aligned}
\mathbb{P}[k_n \leq |\mathcal{C}(v)| \leq \alpha n] &\leq \sum_{t=k_n}^{\alpha n} e^{-n \sup_{s\leq 0}\{s\frac{t}{n} - \log \mathbb{E}[e^{sX}]\}} \\
&\leq \sum_{t=k_n}^{\alpha n} e^{-n \sup_{s\leq 0}\{s\frac{t}{n} - \log\left(e^{-ct/n} + (1-e^{-ct/n})e^{s}\right)\}} \\
&\leq \sum_{t=k_n}^{\alpha n} e^{-n \sup_{s\leq 0}\{s\frac{t}{n} - \log\left(1 + (1-e^{-ct/n})(e^{s}-1)\right)\}} \\
&\leq \sum_{t=k_n}^{\alpha n} e^{\inf_{s\geq 0}\{st - n\log\left(1 + (1-e^{-ct/n})(e^{-s}-1)\right)\}}
\end{aligned}$$

Using the fact that $\log(1 + x) < e^x$,

$$\mathbb{P}[k_n \leq |\mathcal{C}(v)| \leq \alpha n] \leq \sum_{t=k_n}^{\alpha n} e^{\inf_{s\geq 0}\{st - n\left(1-e^{-ct/n}\right)\left(e^{-s}-1\right)\}} \tag{23}$$

By simple derivation, the minimum of the function $s \mapsto st + n(1 - e^{-ct/n})(e^{-s} - 1)$ is attained at

$$s = s^* = \log \frac{n(1 - e^{-ct/n})}{t}$$

Write $t = \beta n$ and $g(\beta, c) = (1 - e^{-c\beta})/\beta$ so that $s^* = \log(g(\beta, c))$. By L'Hôpital's rule :

$$\lim_{\beta \to 0^+} g(\beta, c) = \lim_{\beta \to 0^+} ce^{-c\beta} = c \geq 1$$

By definition, we also have $g(\zeta, c) = 1$. Furthermore :

$$\begin{aligned}
\frac{\partial}{\partial \beta} g(\beta, c) &= \frac{1}{\beta^2}\left(\beta c e^{-c\beta} - (1 - e^{-c\beta})\right) \\
&= \frac{e^{-c\beta}}{\beta^2} \underbrace{\left(c\beta - (e^{c\beta} - 1)\right)}_{<0}
\end{aligned}$$

As a result, $s^* \geq 0$ precisely when $t < \zeta n$, and equation (23) yields

$$\begin{aligned}\mathbb{P}[k_n \leq |\mathcal{C}(v)| \leq \alpha n] &\leq \sum_{t=k_n}^{\alpha n} e^{-t(g(t/n,c)-1-\log g(t/n,c))} \\ &\leq \sum_{t=k_n}^{\alpha n} e^{-t\mathcal{I}_{g(t/n,c)}}\end{aligned}$$

With $\mathcal{I}_c = c - 1 - \log c$

Since $c \mapsto \mathcal{I}_c$ is increasing and $\beta \mapsto g(\beta, c)$ is decreasing, and since $t/n \leq \alpha < \zeta$, it follows that

$$\mathbb{P}[k_n \leq |\mathcal{C}(v)| \leq \alpha n] \leq \sum_{t=k_n}^{\alpha n} e^{-t\mathcal{I}_{g(\alpha,c)}}$$

Denoting $J(\alpha, c) = \mathcal{I}_{g(t/n,c)}$:

$$\mathbb{P}[k_n \leq |\mathcal{C}(v)| \leq \alpha n] \leq \sum_{t=k_n}^{\alpha n} e^{-tJ(\alpha,c)} \leq \frac{e^{-k_n J(\alpha,c)}}{1 - e^{-J(\alpha,c)}}$$

Finally, let $\theta = KJ(\alpha, c) - 1$ so that θ is strictly positive for sufficiently large K. With $C = \frac{1}{1-e^{-J(\alpha,c)}} < \infty$, the probability to have one component with size between k_n and αn is

$$\begin{aligned}\mathbb{P}[\exists v,\ k_n \leq |\mathcal{C}(v)| \leq \alpha n] &\leq n\frac{e^{-k_n J(\alpha,c)}}{1 - e^{-J(\alpha,c)}} \\ &\leq Cne^{-K(\ln n)J(\alpha,c)} \\ &= O(n^{-\theta})\end{aligned}$$

This is true for any $\theta > 0$, taking a constant K sufficiently large. □

Using a similar approach, we can prove that for any $\alpha = \zeta + \delta > \zeta$

$$\mathbb{P}[\exists v, |\mathcal{C}(v)| \geq \alpha n] = O(n^{-\theta})$$

Therefore, for sufficiently large K, with probability at least $(1 - n^{-\theta})$, for any arbitrarily large θ, there is no connected awkward component.

Escape probability Let α be the probability that the total progeny is not small,

$$\alpha = \mathbb{P}[|\mathcal{C}(v)| > S]$$

This is defined as the *Escape probability*, i.e. our branching process escaped an early death. Using the upper and lower bounds seen in Theorems 3.7 and 3.9, we know that

$$\mathbb{P}\left[T^{Bin}_{n-S,p} \geq S\right] \leq \alpha \leq \mathbb{P}\left[T^{Bin}_{n,p} \geq S\right]$$

Using Poisson approximation, both upper and lower bounds of α are asymptotic to $\mathbb{P}\left[T^{Po}_c \geq S\right]$. Using the fact that c is fixed and that $S = K \ln n \to \infty$ when $n \to \infty$:

$$\alpha \sim \mathbb{P}\left[T^{Po}_c \geq S\right] \sim \mathbb{P}\left[T^{Po}_c = \infty\right]$$

By definition, ζ is the survival probability of a Poisson branching process with parameter c, $\zeta = \mathbb{P}\left[T_c^{Po} = \infty\right]$, therefore

$$\alpha \sim \zeta$$

Because the probability of having an awkward component is $o(n^{-k})$ for any k, then the probability of having a giant component is also $\sim \zeta$. Each giant component has size between L^- and L^+ hence between $(\zeta - \delta)n$ and $(\zeta + \delta)n$.

Uniqueness by sprinkling Set q such that $n^{-2} \ll q \ll n^{-1}$, e.g. $q = n^{-3/2}$. Recall that $p = c/n$ so that $q \ll p$. Let $Q \sim G(n, q)$ selected independently from $P \sim G(n, p)$ on the same vertex set, and let $G = P \cup Q$, so that $G \sim G(n, r)$ with $r = p + q - pq$. Indeed, for all edge e,

$$\begin{aligned}\mathbb{P}[e \in G] &= 1 - \mathbb{P}[e \notin G] \\ &= 1 - (\mathbb{P}[e \notin P] \cdot \mathbb{P}[e \notin Q]) \\ &= 1 - (1-p)(1-q) \\ &= p + q - pq\end{aligned}$$

This is often called *sprinkling* as from $G(n, p)$ we sprinkle additional edges coming from $G(n, q)$ with $q \ll p$.

Suppose that P has two giant components, on vertex sets V_1 and V_2. Given that each one has size greater than $L^- = (\zeta - \delta)n$, there are $\Omega(n^2)$ pairs $\{v_1, v_2\}$ with $v_1 \in V_1$ and $v_2 \in V_2$. Because $q \gg n^{-2}$, then a.a.s at least one of these pairs is in the graph Q. Therefore in the graph $G = P \cup Q$, there is a component merging V_1 and V_2, hence of size at least $2(\zeta - \delta)n$.

However by definition $q \ll n^{-1}$, therefore $r \sim p = c/n$ and this component is awkward in G. Therefore P has more than one giant component with probability $o(n^{-\theta})$.

Summary By making δ arbitrarily small, $G(n, p)$ has an expected number $\sim \zeta n$ of points in giant components, and giant component all have size $\sim \zeta n$. By sprinkling argument, the contribution from the possibility that $G(n, p)$ has more than one giant component is negligible : With probability $1 - o(1)$ there is exactly one giant component of size $\sim \zeta n$.

Finally using the result on awkward components, we also know the asymptotic size of the second largest component : $L_2 \leq S = O(\ln n)$.

3.6.2 Barely supercritical phase

Let $p = (1 + \varepsilon)/n$ with $\varepsilon = \lambda n^{-1/3}$ and $\lambda \to \infty$. Note that $\varepsilon^{-2} = \lambda^{-2} n^{2/3} \ll 2\varepsilon n$. The analysis of the barely supercritical phase becomes more difficult as $\lambda(n) \to \infty$ more slowly. We shall then make the assumption that that $\lambda \gg \ln n$.

The following results are taken from [AS92]. Given the length of some part of the proofs, we will only give here quick summaries. Additional details can be found in Łuczak [Łu90], showing that if $\lambda \to \infty$ then a dominant component exists. The *dominant* definition is equivalent to the giant, with relaxed values.

Set $S = K\varepsilon^{-2} \ln n$, $L^- = (1 - \delta)2\varepsilon n$ and $L^+ = (1 + \delta)2\varepsilon n$. Again, a component $\mathcal{C}(v)$ is small if its size is less than S, dominant if $L^- < \mathcal{C}(v) < L^+$, and awkward otherwise. We have similar results than in the very supercritical regime:

- No middle ground : The probability of having any awkward component is $o(n^{-\theta})$ with θ arbitrarily large by taking L arbitrarily large.

  Again we bound the probability that a connected component has size t using bounds on $\mathbb{P}[Bin(n-1, 1-(1-p)^t) = t-1]$. Denoting by μ and σ^2 the mean and variance of this binomial, in the range $t \in [S, L^-]$, we have $\sigma^2 \sim \mu$, and then one can show that, if $t = o(n\varepsilon)$ or $t \sim xn\varepsilon$ with $x \neq 2$, then $\sigma \sim t$ and we can bound the probability by $exp\left(-\Omega((n\varepsilon)^2/t)\right)$.

- Escape probability : Let α be the probability that the total progeny is not small,

  $$\alpha = \mathbb{P}[|\mathcal{C}(v)| > S]$$

  With probability $\alpha \sim 2\varepsilon$ there is a dominant component.

  Again using the upper and lower bounds seen in Theorems 3.7 and 3.9, we know that

  $$\mathbb{P}\left[T^{Bin}_{n-S,p} \geq S\right] \leq \alpha \leq \mathbb{P}\left[T^{Bin}_{n,p} \geq S\right]$$

  And we can show that

  $$\alpha \sim 2\varepsilon$$

  Because there is no middle ground, not small is equivalent to dominant. $C(v)$ is dominant with probability 2ε, and each dominant component has size between $(1-\delta)2n\varepsilon$ and $(1+\delta)2n\varepsilon$.

- By using sprinkling : $L_1 \sim 2n\varepsilon$ and $L_2 \leq K\varepsilon^{-2} \ln n$

  Let $q = n^{-4/3}$ and $Q \sim G(n, q)$ on the same vertex set than $P \sim G(n, p)$. We define $G = P \cup Q$ so that $G \sim G(n, r)$ with $r = p + q - pq = 1 + \varepsilon + o(\varepsilon)$.

  Suppose that P has two dominant component, with vertex sets V_1 and V_2. There are $\gg n^{4/3}$ pairs $\{v_1, v_2\}$, hence with high probability at least one of these pairs is in Q. Therefore G has a connected component of size at least $(1-\delta)4n\varepsilon$, awkward in $G(n, r)$. As P has two dominant components with probability $o(n^{-\theta})$, this proves that

  $$\boxed{L_1 \sim 2n\varepsilon} \qquad \text{and} \qquad \boxed{L_2 \leq K\varepsilon^{-2} \ln n}$$

3.7 Critical window

Let $p = 1/n + \lambda n^{-4/3}$, for any fixed λ. We state here only an overview of the critical window regime, and refer to [Łu90] and [JLR11] for detailed results.

The critical window regime studies the emergence of large components from trees. We consider the evolution of $G(n, p)$ when λ varies in $\mathbb{R}$. The main result of the critical window states that for a given λ, the sizes of the largest components are of the form $cn^{2/3}$ with a distribution over the constant c with support in $\mathbb{R}^+$. Hence, for any λ and any $c > 0$, asymptotically, with strictly positive probability, the largest component is bigger (resp. smaller) than $cn^{2/3}$.

When λ is largely negative, then the largest component will likely be of the size $\varepsilon n^{2/3}$ with ε small, many components will have about the same size, and most of the components are trees. When λ grows, the dominant component starts to emerge, of non-simple complexity, and the second largest component is simple and far smaller.

Suppose now that for a given λ, a given Graph in $G(n,p)$ has components of size $c_1 n^{2/3}$ and $c_2 n^{2/3}$. As λ increases to $\lambda + d\lambda$, they will merge with probability $c_1 c_2 d\lambda$, i.e. components merge with a probability proportional to their size.

Finally, with probability $(c^2/2)d\lambda$ a new edge appears in a component of size $cn^{2/3}$, i.e. large components rarely remain trees.

3.8 Discrete Duality Principle

Theorem 3.11. *Let $\mu < 1 < \lambda$ be conjugates, as per the discrete duality principle for Poisson branching process 2.10. The graph $G(n,p)$with $p = \lambda/n$ after removing the giant component is close in law to the random graph $G(m,p')$ with $p' = \mu/m$ and where $m = \lceil n(1-\zeta)\rceil$ is the asymptotic number of vertices outside the giant component (recall that ζ is the survival probability of our graph branching process).*

Proof. Let $p = \lambda/n$ with $\lambda > 1$. Note first that, conditionally on not being in the giant component, all the edges in the complement of the giant component of $G(n,p)$ are independent (by definition of $G(n,p)$). Furthermore, the probability of existence of any edge e in $G(n,p)$ is equal to λ/n.

Let $L_1 = n - m \approx \zeta n$, so that $m \approx \eta n$. The conditional probability that an edge is present in $G(n,p)$ with the giant component removed, conditionally on $L_1 \approx \zeta n$, is equal to

$$\frac{\lambda}{n} = \frac{\lambda}{m}\frac{m}{n} \approx \frac{\lambda\eta}{m}$$

As μ and λ are conjugate, $\mu = \lambda\eta$ and

$$\frac{\lambda}{n} \approx \frac{\mu}{m}$$

□

4. Intermed - Given degree sequences

As explained in the introduction, the traditional Erdős-Rényi model of a random network is of little use in modelling the type of complex networks which researchers study nowadays. Modern networks are of diverse nature and usually exhibit inhomogeneity among their nodes and correlations among their edges. For example, we observe empirically in the web that certain important pages will have many more links entering them than typical ones. This motivates the study, for a fixed degree sequence $\mathcal{D} = (d_1, \dots, d_n)$, of graphs $G(\mathcal{D})$ on n vertices where vertex i has degree d_i.

Ideally we would like to investigate *uniform graphs* having a given degree sequence. It is not a trivial task to generate graphs having a specific degree sequence, mainly because they may not exists : For a given $\mathcal{D} = (d_1, \dots, d_n)$, the set of undirected simple graphs with n vertices where vertex j has degree d_j may be empty. First, for such a graph to exist, we must have an even total degree (as per by hand-shaking lemma, this is equal to twice the number of edges),

$$\ell_n = \sum_{j\in[n]} d_j \in 2\mathbb{Z} \tag{24}$$

However this is not a sufficient condition. Consider for instance the graph on $n = 4$ vertices with degree sequence $(3,3,1,1)$. The two first vertices are full-degree, they are both connected to all other vertices,

so that the minimum degree of such a graph is 2. Therefore there is no simple graph on 4 vertices with such a degree sequence. This feasibility question has been settled by Erdős and Gallai. See appendix A.4 for some details.

We say that a degree sequence $\mathcal{D}$ is *graphic* if $\mathbb{G}(\mathcal{D})$, the set of graphs G on n vertices such that for all $i \in [n]$, $deg(i, G) = d_i$, is not empty. If $\mathcal{D}$ is graphic, we may then define $G(\mathcal{D})$ as the uniform probability distribution on $\mathbb{G}(\mathcal{D})$. We will often abuse the notation, with $G(\mathcal{D})$ being a graph on n vertices chosen uniformly at random among all graph with degree sequence $\mathcal{D}$.

The study of the existence of a giant component in random graphs with an arbitrary given degree sequence started with the result of Molloy and Reed [MR95]. Although they define the concept of asymptotic degree sequences, in order to be consistent with later results from Joos, Perarnau, Rautenbach and Reed [JPRR18], we will rather define sequences of degree sequences $\mathfrak{D} = (\mathcal{D}_n)_{n\geq 1}$. Using a symmetry argument, one can easily translate results for sequences of degree sequences to asymptotic degree sequences, and vice versa.

For every $\mathcal{D}_n = (d_1^{(n)}, \dots, d_n^{(n)})$, we define $n_i = n_i(n) = |\{j \in [n] : d_j^{(n)} = i\}|$, the number of vertices with degree i in the degree sequence $\mathcal{D}_n$ of length n.

Definition 4.1. Let $\mathfrak{D}$ be a sequence of degree sequence. We say that $\mathfrak{D}$ is

- **feasible**, if for every $n \geq 1$, there exists at least one simple graph on n vertices with degree sequence $\mathcal{D}_n$. (i.e. every degree sequence $\mathcal{D}_n$ is graphic).
- **smooth** if for every integer $i \geq 0$, there exists $\lambda_i \in [0, 1]$ such that $\lim_{n\to\infty} \frac{n_i}{n} = \lambda_i$.
- **sparse** if there exists $\lambda \in (0, \infty)$ such that $\lim_{n\to\infty} \sum_{i\geq 1} \frac{in_i}{n} = \lambda$ (i.e. the average degree is linear in n).
- **f-bounded** for some function $f : \mathbb{N} \to \mathbb{R}$, if $n_i = 0$ for every $i > f(n)$. There are no vertices of degree higher than $f(n)$.

We only consider degree sequences such that $n_0 = 0$ as isolated vertices play no role in the composition of any component. Given a smooth sequence of degree sequence $\mathfrak{D}$, we define the following key parameters:

$$\begin{cases} K & = \sum_i i\lambda_i \\ \mathcal{Q}(\mathfrak{D}) & = \sum_i i(i-2)\lambda_i \end{cases}$$

Note that K is a measure of the average degree in our graph.

Following the denomination in [JPRR18], we say that $\mathfrak{D}$ satisfies the **MR**-conditions if

MR1. it is feasible, smooth and sparse,

MR2. it is f-bounded by the function $f : n \mapsto n^{1/4-\varepsilon}$, for some $\varepsilon > 0$,

MR3. for every $i \geq 1$, $\frac{i(i-2)n_i}{n}$ converges uniformly to $i(i-2)\lambda_i$, i.e for all $\varepsilon > 0$ there exist N such that for all $n > N$ and for all $i \geq 0$,

$$\left| \frac{i(i-2)n_i}{n} - i(i-2)\lambda_i \right| < \varepsilon$$

MR4. $L(\mathfrak{D}) = \lim_{n\to\infty} \sum_{i\geq 1} i(i-2)\frac{n_i}{n}$ exists and converges uniformly to $\sum_{i\geq 1} i(i-2)\lambda_i$, i.e.

(a) if $L(\mathfrak{D})$ is finite then for all $\varepsilon > 0$ there exists i^* and N such that for all $n > N$

$$\left| \sum_{i=1}^{i^*} i(i-2)n_i/n - L(\mathfrak{D}) \right| < \varepsilon$$

(b) if $L(\mathfrak{D})$ is infinite then for all $T > 0$, there exists i^* and N such that for ll $n > N$

$$\sum_{i=1}^{i^*} i(i-2)n_i/n > T$$

5. Configuration model

5.1 Presentation of the model

A major difficulty in the study of random graphs on fixed degree sequences is that it is difficult to generate such graphs directly. Instead it has become standard to study random configurations on a fixed degree sequence, and use some lemmas to translate results from random configurations to random graphs. The configuration model was introduced by Bender and Canfield [BC78] and refined by Bollobas [Bol80] and Wormald [Wor80].

Let $\mathcal{D} = (d_1, \dots, d_n)$ be a degree sequence such that $\sum_{i=1}^{n} d_i$ is even. Then there exists multigraphs with degree sequence $\mathcal{D}$. Instead of building a probability distribution on $\mathbb{G}(\mathcal{D})$, we will do so on $\widehat{\mathbb{G}}(\mathcal{D})$, the set of multigraphs on n vertices such that for all $i \in [n]$, $deg(i, G) = d_i$.

First construction on half-edges

The first construction for a (multi)graph based on a configuration used the notion of half-edges.

Let $\Delta = \{(i, j),\ 1 \leq i \leq n,\ 1 \leq j \leq d_i\}$. It can be visualized as a set of half-edges. For instance let $\mathcal{D} = (2, 2, 1, 3)$, then we have Δ:

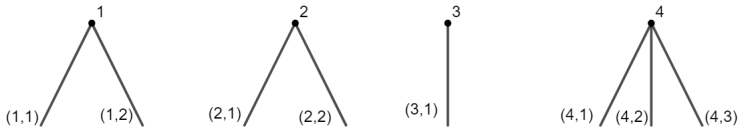


Figure 6: Set Δ

Our aim is to match each half-edge with another one, forming a multigraph (with possible loops and multi-edges). To do so, we define $M(\Delta)$ to be the set of involutions acting on $[n]$ (i.e. permutation such that $\sigma^{-1} = \sigma$), with no fixed point (derangements). $M(\Delta)$ is called the set of *matching* of Δ.

Note that if $|\Delta|$ is even, then by simple combinatoric argument we know the cardinality of $M(\Delta)$:

$$|M(\Delta)| = (|\Delta| - 1) \cdot (|\Delta| - 3) \cdot \ldots \cdot 3 \cdot 1$$

For example, keeping the same Δ, figure 7, represents one possible matching for Δ.

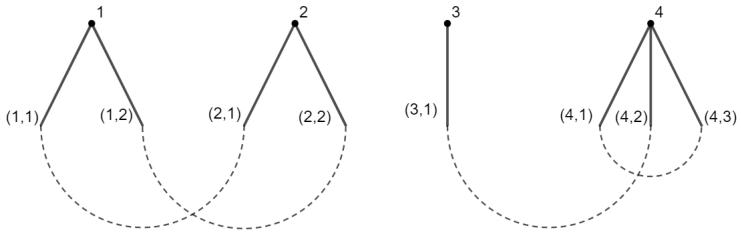


Figure 7: A matching on the set Δ

For any matching $\sigma \in M(\Delta)$, we then build a multigraph on $[n]$ with edge set:

$$E = \{(i, i'),\ \sigma(i,j) = (i', j'),\ (i,j) \in \Delta\}$$

Figure 8 illustrates the graph obtained from the matching in figure 7.

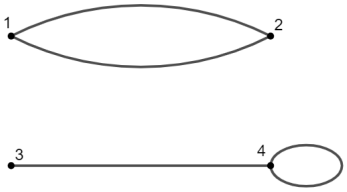


Figure 8: The underlying multigraph

Let σ be a random matching of Δ drawn uniformly among all matchings. Then, we may define the random multigraph $G = G(\sigma)$ on $[n]$. We denote by $\widehat{\mathcal{G}}(\mathcal{D})$ the corresponding probability distribution on $\widehat{\mathbb{G}}(\mathcal{D})$. This is the **configuration model**.

Given a degree sequence $\mathcal{D}$, a *configuration* is then a pair (Δ, σ) with $\sigma \in M(\Delta)$.

Equivalent construction on copies of vertices

In the following, we will use an alternative equivalent construction for a (multi)graph from a configuration, using copies of vertices.

Recall that we have the degree sequence $\mathcal{D} = (d_1, \dots, d_n)$. Let define the set L as containing $deg(v)$ distinct copies of each vertex v. Then with $M(L)$ the set of matching of L, we can select a random matching in $M(L)$.

With our example $\mathcal{D} = (2, 2, 1, 3)$,

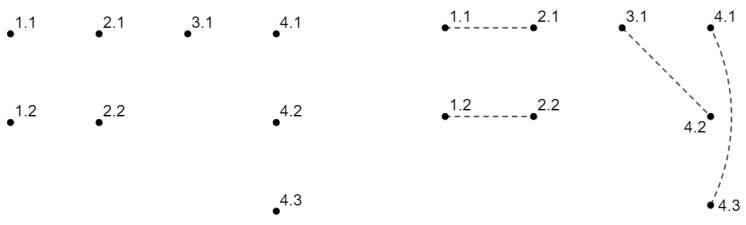


Figure 9: The set L and a matching on the set L

A configuration is then a pair (L, σ), with $\sigma \in M(L)$. It defines a multigraph $G(\sigma)$ whose edges are the pairs in the matching. Abusing the language, we'll say that a configuration has a graphical property P (e.g. is planar) if the underlying multigraph does. Figure 10 illustrates the graph obtained from the matching set on L as seen in 9.

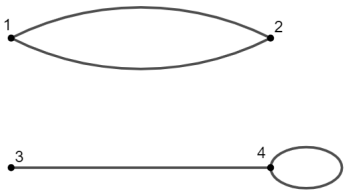


Figure 10: The underlying multigraph $G(\sigma)$

From configuration to graphs

In [McK85], McKay gave an upper bound for the number of labelled graphs with a given degree sequence. From the main result in his article, we can deduce the following lemmas, allowing us to prove results about a graph by proving the statement for a configuration.

Lemma 5.1. *Let $\mathfrak{D} = (\mathcal{D}_n)_{n\geq 1}$ be a sequence of degree sequence that satisfies the **MR**-conditions with $\mathcal{Q}(\mathfrak{D}) < \infty$. If a random configuration with degree sequence $\mathcal{D}_n$ a.s. has a property P, then a random graph with the same degree sequence a.s. has property P*

Lemma 5.2. *Let $\mathfrak{D} = (\mathcal{D}_n)_{n\geq 1}$ be a sequence of degree sequence that satisfies the **MR**-conditions (with $\mathcal{Q}(\mathfrak{D})$ possibly unbounded). If a random configuration with degree sequence $\mathcal{D}_n$ has a property P with probability at least $1 - z^n$ for some constant $0 < z < 1$, then a random graph with the same degree sequence a.s. has property P*

His exact result is the following. Given a simple graph X on n vertices with max degree x_{max}, and given a degree sequence $\mathcal{D} = (d_1, \dots, d_n)$ with maximum degree d_{max}, define

$$\lambda = \frac{1}{4m}\sum_{i=1}^{n} d_i^2 \qquad \text{and} \qquad \mu = \frac{1}{2m}\sum_{v_i v_j \in E(X)} d_i d_j$$

with $m = \sum_{i=1}^{n} d_i$. Then

Theorem 5.3 (Brendan D. McKay [McK85])**.** *Suppose that $d_{max} \geq 1$, and $\widehat{\Delta} \leq \varepsilon_1 m$ where $\varepsilon_1 < 2/3$ and $\widehat{\Delta} = 2 + d_{max}\left(\frac{3}{2}d_{max} + x_{max} + 1\right)$. Then the number of simple graph, with degree sequence $\mathcal{D}$ and no edges in common with X, is uniformly*

$$\frac{(2m)!}{m!2^m \prod_{i=1}^{n} d_i!} exp\left(-\lambda - \lambda^2 - \mu + O(\widehat{\Delta}^2/m)\right)$$

Let $\mathcal{M}(\mathcal{D})$ be the set of symmetric $0-1$ matrices of order n, with zero diagonal elements and row sums $d_1, \dots, d_n$ respectively, and call $\mathcal{N}(\mathcal{D})$ the cardinality of $\mathcal{M}(\mathcal{D})$. Define $P(\mathcal{D})$ as

$$P(\mathcal{D}) = \mathcal{N}(\mathcal{D})\frac{m!2^m \prod_{i=1}^{n} d_i!}{(2m)!}$$

We know that the number of possible matching in $M(L)$ is exactly $|M(L)| = \frac{(2m)!}{m!2^m}$ and that each matrix in $\mathcal{M}(\mathcal{D})$ corresponds to exactly $\prod_{i=1}^{n} d_i!$ possible matching. Therefore $P(\mathcal{D})$ can be understood as the probability that a random pairing in $M(L)$ induces a simple graph (with neither loops nor multiple edges). The result then follow from estimating $P(\mathcal{D})$.

5.2 Exploration process

In order to study the different components of a random configuration, we need to define an exploration process.

Given a sequence of degree sequences $\mathcal{D}_n = (d_1^{(n)}, \dots, d_n^{(n)})$, we will explore a random configuration F on n vertices, with n_i of degree i. We will use the following terminology,

Entirely exposed A vertex all of whose copies are in exposed pairs.

Partially exposed A vertex some but not all of whose copies are in exposed pairs.

Unexposed All other vertices are unexposed.

Open The copies of partially exposed vertices which are not in explored pairs.

The exploration process runs as follow, from the set L seen above (see section 5.1),

(1) Expose a pair of the configuration by first choosing any member of L, and then choosing its partner at random. Remove the pair from L.

(2) If there exists at least one partially exposed vertex, Choose a random open copy of a partially exposed vertex, and pair it with another random member of L. Remove the pair from L. Repeat until there is no partially exposed vertex.

(3) Repeat step (1) and (2) until L is empty.

All random choices are made uniformly.

Step (1) corresponds to the start of a new connected component in our graph, while (2) is repeated until one component is fully explored.

Now, let X_i be the number of open copies after the i^{th} step in the exploration process. At each step (1), X_i is set at $d_1 + d_2 - 2$ if the copies chosen are from vertices of degree d_1 and d_2. Then at each occurrence of step (2), if we neglect (for now) the case where the pair is taken from two open copies, X_i increases by $d - 2$ where d is the degree of the vertices v whose copy has been chosen to pair our open copy. Therefore, initially, the expected increase of X_i is

$$\begin{aligned}
\mathbb{E}[\Delta X_i] &= \sum_{d \geq 1} (d-2)\mathbb{P}[v \text{ has degree } d] \\
&= \sum_{d \geq 1} (d-2)\frac{d n_d}{\sum_{i \geq 1} i n_i} \\
&= \frac{\mathcal{Q}(\mathfrak{D})}{K}
\end{aligned}$$

This give a rationale for the study of $\mathcal{Q}(\mathfrak{D})$ in our giant component analysis. If it is negative, then X_i promptly goes to 0, the component is explored, and X_i resets itself by exploring a new component. But if $\mathcal{Q}(\mathfrak{D})$ is positive, then initially the expected increase of X_i is positive and we will prove that a.s. our random configuration has a component of size $\Theta(n)$.

Note however that as exploration process advances, the expected increase of X_i shifts from this initial values, impacted by the fact that the ratio of members of L which are open copies of vertices of degree d changes.

In addition to X_i the number of open vertex-copies after i pairs of our configuration have been exposed, we also introduce the following variables:

(1) Y_i is the number of *backedges* formed after i steps. A backedge is an edge formed between two open vertex-copies. In this case (neglected in the above approximation of the initial expected increase of X_i), X_i decreases by 2.

(2) C_i is the number of components that have been at least partially exposed.

(3) W_i is the sum $\sum_v(deg(v)-2)$ over all vertices completely or partially exposed during the first i steps. We note that
$$W_i = X_i + 2Y_i - 2C_i$$
Note that W_i remains unchanged whenever a backedge is created. It is then easier to analyse W_i indexed by the number of new vertices exposed rather than by the number of pairs exposed. This is the following variable,

(4) Z_j, the sum $\sum_v(deg(v)-2)$ over the first j new vertices completely or partially exposed. Remark that Z_j has the same initial expected increase than X_i.

(5) In order to compare Z_j and W_i, we introduce I_j to be the number of pairs exposed by the time that the j^{th} vertex is partially exposed, i.e.
$$W_{I_j} = Z_j$$

Finally, in order to get a precise estimation of the giant component, we will also follow the degree of the unexposed vertices at each time step. We define $d_{i,j}$ as the number of unexposed vertices of degree i after the exposition of j pairs of the configuration. Hence $d_{i,0} = n_i \approx \lambda_i n$.

5.3 Statement of results

We recall our definition for a degree sequence $\mathfrak{D}$ satisfying the **MR**-conditions:

(1) it is feasible, smooth and sparse,

(2) it is f-bounded by the function $f : n \mapsto n^{1/4-\varepsilon}$, for some $\varepsilon > 0$,

(3) for every $i \geq 1$, $\frac{i(i-2)n_i}{n}$ converges uniformly to $i(i-2)\lambda_i$,

(4) $L(\mathfrak{D}) = \lim_{n\to\infty} \sum_{i\geq 1} i(i-2)\frac{n_i}{n}$ exists and converges uniformly to $\sum_{i\geq 1} i(i-2)\lambda_i$.

Theorem 5.4 (Molloy and Reed [MR95]). *Let $\mathfrak{D} = (\mathcal{D}_n)_{n\geq 1}$ be a sequence of degree sequence that satisfies the **MR**-conditions, and let $G(\mathcal{D}_n)$ be a graph on n vertices chosen uniformly at random among all graph with degree sequence $\mathcal{D}_n$. Then,*

(a) *if* $\mathcal{Q}(\mathfrak{D}) < 0$ *and the sequence is* $n^{1/8-\varepsilon}$*-bounded for some* $\varepsilon > 0$*, then for every constant* $\zeta > 0$ *the probability that* $G(\mathcal{D}_n)$ *has no component of order at least* ζn *is* $1 - o(1)$*.*

(b) *if* $\mathcal{Q}(\mathfrak{D}) > 0$ *then there exists a constant* $\zeta > 0$ *such that the probability that* $G(\mathcal{D}_n)$ *has a component of order at least* ζn *is* $1 - o(1)$*. Furthermore if* $\mathcal{Q}(\mathfrak{D})$ *is finite, then* $G(\mathcal{D}_n)$ *has exactly one component of order greater than* $\gamma \log n$ *for some constant* γ *dependent on* $\mathcal{D}_n$*.*

Furthermore, define the function $\chi : [0, 1] \to \mathbb{R}$ by:

$$\alpha \mapsto K - 2\alpha - \sum_{i \geq 1} i\lambda_i \left(1 - \frac{2\alpha}{K}\right)^{\frac{i}{2}} \tag{25}$$

If the equation $\chi(\alpha) = 0$ admits a solution, let $\alpha_{\mathfrak{D}}$ be the smallest positive solution, and let

$$\zeta_{\mathfrak{D}} = 1 - \sum_{i \geq 1} \lambda_i \left(1 - \frac{2\alpha_{\mathfrak{D}}}{K}\right)^{\frac{i}{2}} \tag{26}$$

$$\lambda'_i = \frac{\lambda_i}{1 - \zeta_{\mathfrak{D}}} \left(1 - \frac{2\alpha_{\mathfrak{D}}}{K}\right)^{\frac{i}{2}} \tag{27}$$

Then, in [MR98], Molloy and Reed gave a more precise estimation on the size of the giant component when $\mathcal{Q}(\mathfrak{D}) > 0$.

Theorem 5.5 (Molloy and Reed [MR98])**.** *Let* $\mathfrak{D} = (\mathcal{D}_n)_{n \geq 1}$ *be a sequence of degree sequence that satisfies the* ***MR****-conditions and such that* $\mathcal{Q}(\mathfrak{D}) > 0$*, and let* $G = G(\mathcal{D}_n)$ *be a graph on* n *vertices chosen uniformly at random among all graph with degree sequence* $\mathcal{D}_n$*. Then, in addition to theorem 5.4, if* $\mathcal{Q}(\mathfrak{D}) > 0$*, the giant component of* G *has order* $\zeta_{\mathfrak{D}} n + o(n)$*.*

This allow for a discrete duality principle in the configuration model,

Theorem 5.6 (Molloy and Reed [MR98])**.** *Let* $\mathfrak{D} = (\mathcal{D}_n)_{n \geq 1}$ *be a sequence of degree sequence that satisfies the* ***MR****-conditions and such that* $\mathcal{Q}(\mathfrak{D}) > 0$*, and let* $G = G(\mathcal{D}_n)$ *be a graph on* n *vertices chosen uniformly at random among all graph with degree sequence* $\mathcal{D}_n$*. Then, almost surely, the structure of the graph formed by deleting the largest component* $\mathcal{C}$ *from* G *is essentially the same as that of a random graph on* n' *vertices with*

$$n' = n - |\mathcal{C}| = (1 - \zeta_{\mathfrak{D}})n + o(n)$$

with degree sequence $\mathcal{D}'_n = (d_1^{'(n)}, \dots, d_n^{'(n)})$ *such that* $n'_i = \lambda'_i n + o(n)$*.*

5.4 Proof of Theorem 5.4 - Existence of a giant component

We will only give here an outline of the proof of Theorem 5.4. We refer the reader to [MR95] for an extended version.

5.4.1 Subcritical phase - graphs with no large components

We place ourselves in the subcritical case $\mathcal{Q}(\mathfrak{D}) < 0$. Let $\nu = -\mathcal{Q}(\mathfrak{D})/K$ and set $R = 150/\nu^2$. We start by stating an equivalent of Theorem 5.4 for configurations,

Theorem 5.7. *Let F be a random configuration with n vertices and degree sequence $\mathfrak{D}$, meeting the conditions of Theorem 5.4. If $\mathcal{Q}(\mathfrak{D}) < 0$ and if, for some positive function $\phi(n) \leq n^{1/8-\varepsilon}$, F has no vertices of degree greater than $\phi(n)$, then F has a.s. no components with more than $\alpha = \lceil \phi(n)^2 R \log n \rceil$ vertices.*

Lemma 5.8. *Suppose that F is as described in Theorem 5.7. Given any vertex v in F, the probability that v lies in the component of size at least α is less then n^{-2}.*

$$\mathbb{P}[\{v \in C,\ |C| \geq \alpha\}] < n^{-2}$$

Proof. Note that w.l.o.g, we can assume that v is the first vertex chosen in step (1) of our exploration process, i.e. the first vertex chosen when starting the exploration of the first connected component. This is because we have complete freedom as to which vertex we pick to start the exploration. Therefore if the results hold for this first vertex, it will hold for any vertex from F, and the probability that v belongs to a component of size at least α is at most the probability that $X_i > 0$ for all $1 \leq i \leq \alpha$.

Then using that any such i we have $C_i = 1$, one can show that:

$$W_i \leq X_i - 2 \leq Z_i \tag{28}$$

The probability that $X_i > 0$ for all $1 \leq i \leq \alpha$ is at most the probability that $Z_\alpha > -2$ (using that Z_i is decreasing on $i \in \{1 \dots \alpha\}$).

The initial expected increase of Z_i is

$$\mathbb{E}[Z_1] = \frac{\sum_{i\geq 1} i(i-2)n_i}{\sum_{i\geq 1} i n_i} = \frac{\mathcal{Q}(\mathfrak{D})}{K} + o(1) = -\nu + o(1)$$

Now for any $j \leq \alpha$, the expected increase of Z_j is maximized when the first j vertex-copies chosen are all copies of vertices of degree 1. In this case we would have,

$$\mathbb{E}[Z_{j+1} - Z_j] = \frac{-(n_1 - j) + \sum_{i\geq 2} i(i-2)n_i}{(n_1 - j) + \sum_{i\geq 2} i n_i} + o(1)$$

Using that $\alpha = o(n)$, then for any $j \leq \alpha$, $j = o(n)$ and because $in_i \to \lambda_i$ uniformly,

$$\mathbb{E}[Z_{j+1} - Z_j] = -\nu + o(1) \leq -\frac{\nu}{2}$$

Therefore,

$$\mathbb{E}[Z_\alpha] \leq -\frac{\nu}{2}\alpha + deg(v) < -\frac{\nu}{3}\alpha = \Gamma$$

Recall the following corollary of Azuma's inequality (see A.5),

Corollary. *Let $\Sigma = \Sigma_1, \Sigma_2, \dots, \Sigma_n$ be a sequence of random events. Let $f(\Sigma)$ be a random variable defined by these Σ_i. If for each i,*

$$\max \left| \mathbb{E}\left[f(\Sigma) \mid \Sigma_1, \Sigma_2, \dots, \Sigma_{i+1}\right] - \mathbb{E}\left[f(\Sigma) \mid \Sigma_1, \Sigma_2, \dots, \Sigma_i\right] \right| \leq c_i,$$

then

$$\mathbb{P}\left[|f - \mathbb{E}[f]| > t\right] \leq 2e^{\frac{-t^2}{2\sum c_i^2}}.$$

Then, we define Σ_i as the choice of i^{th} new vertex exposed for $i = 1, \dots, \alpha$, and let $f(\Sigma) = Z_\alpha$. We know that the probability that $X_i > 0$ for all $1 \leq i \leq \alpha$ is at most the probability that $Z_\alpha > -2$, therefore using the fact that $\Gamma = \mathbb{E}[Z_\alpha]$ is negative,

$$\mathbb{P}[\{|C| \geq \alpha\} \leq \mathbb{P}[|Z_\alpha - \Gamma| > -\Gamma]$$

We can use this corollary to prove the probability that v lies in the component of size at least α is at most

$$2\exp\left(-\frac{\Gamma^2}{8\sum_{i=1}^{\alpha}\phi(n)^2}\right) = 2n^{-\frac{\nu^2}{72}R} < n^{-2}$$

□

Theorem 5.7 follow immediately : The expected number of vertices in a component of size at least α is $o(1)$, therefore a.a.s none exist.

Finally, using lemma 5.1, we know that this result holds for random graphs, hence proving Theorem 5.4a.

5.4.2 Supercritical phase - graphs with a giant component

We place ourselves in the supercritical case $\mathcal{Q}(\mathfrak{D}) > 0$. As for the subcritical phase, we start by stating an equivalent of Theorem 5.4 for configurations,

Theorem 5.9. *Let F be a random configuration with n vertices and degree sequence $\mathfrak{D}$, meeting the conditions of Theorem 5.4. If $\mathcal{Q}(\mathfrak{D}) > 0$ then there exist ζ_1 such that a.s. F has a component with at least $\zeta_1 n$ vertices. Moreover, the probability of the converse is at most z^n, for some fixed $0 < z < 1$.*

The proof of the theorem in the supercritical phase is substantially longer than the subcritical phase. It relies on the following sequence of lemmas.

Lemma 5.10. *There exists $0 < \varepsilon < 1$ and $0 < \Delta < \min\left(\frac{1}{4}, \frac{K}{4}\right)$ such that for all $0 < \delta < \Delta$, almost surely $Z_{\lceil \delta n \rceil} > \delta n$. Moreover, the probability of the converse is at most z_1^n, for some fixed $0 < z_1 < 1$.*

Proof. Initially, the probability that the vertex-copy chosen has degree i is

$$p_i(n) = \frac{in_i}{\sum_{j\geq 1} jn_j} = \frac{i\lambda_i}{K} + o(1)$$

Contrary to the subcritical section, here we have to study this increase after $\Theta(n)$ steps. After this many steps, the ratio of unexposed vertices of different degrees are shifting, changing the expected increase in X_i.

In order to get round this complication, we define and study Z_j^*, a small variation on Z_j, where

- one can find a value i^* such that any time a new vertex of degree $i > i^*$ is explored, instead of increasing Z_j by $(i-2)$, if we decrease it by 1, we still have a positive expected increase.
- one can exhibit a sequence $\phi_1, \dots, \phi_{i^*}$ summing to one, such that for each $1 < i \leq i^*$, the initial probability of choosing a vertex of degree i (equal to $p_i(n)$) is a just little higher than ϕ_i. Hence in Z_j^*, we modify the exploration process so that this initial probability is exactly ϕ_i, and we can find $\phi_1, \dots, \phi_{i^*}$ such that the expected increase of Z_j^* is still positive.

The random variable Z_j^* follows the random walk:

- $Z_0^* = 0$
- $Z_{j+1}^* = Z_j^* + (i-2)$ with probability ϕ_i for $1 \leq i \leq i^*$

For $i = 2, \dots, i^*$, we choose Δ_i such that $\frac{i\lambda_i - \Delta_i}{K} > \phi_i$ and set $\Delta = \min\{\Delta_2, \dots, \Delta_{i^*}, K/4\}$. For any $i = 2, \dots, i^*$, the probability to chose a vertex of degree i is at least $\frac{i\lambda_i - \Delta_i}{K}$, therefore after Δ iterations, for any $i = 2, \dots, i^*$, this probabilities is at least higher that ϕ_i. Therefore for any $0 \leq j \leq \Delta n$. The random variable Z_j majorize Z_j^* and

$$\mathbb{P}[Z_j > R] \geq \mathbb{P}[Z_j^* > R].$$

For some $\varepsilon > 0$ and for large enough n, one can show that the expected increase of Z_j^* at any step is ε, and using some well-know results on random walks (see [Fel] for example), $Z_{\delta n}^*$ is a.s. concentrated around its expected value $\delta\varepsilon n$ □

This lemma shows that almost surely, Z_j grows large. In order to transfer this result into X_i, we will use the variable I_j defined previously* such that $W_{I_j} = Z_j$. We deduce then the following result on X_i,

Lemma 5.11. *There exists $0 < \delta' < \Delta$ such that for all $0 < \delta < \delta'$, almost surely there exists some $1 \leq I \leq I_{\lceil \delta n \rceil}$ such that $X_I > \gamma n$, where $\gamma = \min\left(\frac{\varepsilon\delta}{2}, \frac{1}{4}\right)$. Moreover, the probability of the converse is at most z_2^n, for some $0 < z_2 < 1$ dependent on δ.*

Proof. We refer to [MR95] for a detailed proof. Note that Δ is a defined in the proof of lemma 5.10, and we find $\delta' = \frac{\Delta}{1+\varepsilon}$. □

Now that we know that X_i is a.a.s large (at least as large as $\Theta(n)$), we can prove that there exists a giant component.

Lemma 5.12. *There exists $\zeta > 0$ such that the component being exposed at step $I = I_{\lceil \delta' n \rceil}$ will almost surely have at least ζn vertices. Moreover, the probability of the converse is at most z_3^n, for some fixed $0 < z_3 < 1$.*

Proof. By definition of δ', we know that at step I, there at least $n/5$ unexposed vertices. We form a set β consisting of exactly one copy of each of them.

There are X_I open-vertex copies, and therefore a set χ of size X_I whose elements must be exposed before the component is entirely exposed.

In L, at step I, there are $M - 2I$ vertex-copies available for matching. The exploration process select a uniformly random matching, therefore the expected number of pairs joining one element of χ (open vertex-copies) to one element of β (one copy per unexplored vertex) is at least

$$|\beta| \frac{X_I}{M - 2I} \geq \frac{n}{5} \frac{X_I}{M - 2I}$$

Then by lemma 5.11, we obtain a lower bound of ζn for this value, and it follows from Chernoff's bound that it is a.s. half of their expected values. Therefore the component a.s. has at least ζn vertices. □

*see section 5.2

Theorem 5.9 follows immediately. To conclude we will prove that F a.s. has only one large component.

Lemma 5.13. *If F is a random configuration as per Theorem 5.9, then a.a.s. F has exactly one component on more than $T \log n$ vertices, for some constant T dependent on the degree sequence.*

Proof. Let u and v be two vertices of G. We say that the ordered pair (u, v) has property A if u belongs to a component of of size at least ζn while v is in another component of size at least $T \log n$ for some T.

Suppose that (u, v) has property A. We can decide to start our exploration process at u. By lemma 5.11, we can assume there is some $I \leq I_{\lceil \delta n \rceil}$ such that $X_I > \gamma n$ with $\gamma = \min\{\frac{\varepsilon\delta'}{2}, \frac{\varepsilon\zeta}{2}, \frac{1}{4}\}$. Note that by definition of property A, at step I we must still be exploring the first component C_1 and v is an unexplored vertex. Let χ be the set of open-vertex copies at step I.

We will change here our standard exploration process. Delay the exploration of the remaining of C_1, and start the exploration of v's component C_2. We will see that if C_2 gets too big then a.s it must include an element from χ, hence reaching a contradiction.

Start the exploration of C_2 from v. At each step, the probability to chose a member of χ to partner our open copies of C_2 is at least γ/K. Then, for an appropriate T depending on K, the probability that v lies in a component of size $T \log n$, not intersecting χ is at most

$$\left(1 - \frac{\gamma}{K}\right)^{T \log n} = o(n^{-2})$$

The expected number of pairs of vertices with property A tends to zero as n tends to infinity, therefore a.a.s none exist. □

Finally, lemma 5.2 allows us to transfer the results from Theorem 5.9 and from the uniqueness of the giant component into graphs, proving Theorem 5.4b.

5.5 Proof of Theorem 5.5 - Size of the giant component

The theorem 5.5 is an improvment of theorem 5.4. While the former proved in the supercritical phase (hence when $\mathcal{Q}(\mathfrak{D}) > 0$) the existence of a constant ζ such that a component with order ζn a.a.s exists, the later will give a precise estimation of this ζ. We will make use of the variable $d_{i,j}$, the number of unexposed vertices of degree i after the exposition of j pairs, and of [Wor95], where Wormald prove that under certain conditions, random variables almost surely behave like solutions to a system of differential equations.

We say that a function $f(x_1, \dots, x_n)$ satisfies a *Lipschitz condition* on $D \subseteq \mathbb{R}^n$ if there exist a constant $L > 0$ such that for any $\vec{x}$ and $\vec{y}$ in D,

$$|f(x_1, \dots, x_n) - f(y_1, \dots, y_n)| \leq L \sum_{i=1}^{n} |x_i - y_i|$$

While the original theorem worked with a system of differential equations, we only need a simplified version, dealing with one differential equation.

Theorem 5.14 ([Wor95])**.** *Let $(Y_t)_{t=0}^{m(n)}$ be a sequence of random variables, taking values in $\mathbb{R}^+$, such that $Y_t \leq Cn$ for some constant C, and let H_t be the history of the sequence, i.e. $H_t = (Y_0, Y_1, \dots, Y_t)$. Suppose also that for some function $f : \mathbb{R}^2 \to \mathbb{R}$, the following hold:*

(i) There is a constant C' such that for all $t < m$

$$|Y_{t+1} - Y_t| < C'$$

(ii) Uniformly over all $t < m$,

$$\mathbb{E}[Y_{t-1} - Y_t \mid H_t] = f(t/n, Y_t/n) + o(1)$$

(iii) the function f is continuous and satisfies a Lipschitz condition on some bounded connected open set D, containing the intersection of $\{(t, z) : t \geq 0\}$ with a neighbourhood of $\{(0, z) : \mathbb{P}[Y_0 = zn] \neq 0$ for some $n\}$.

Then

(1) for $(0, \widehat{z}) \in D$, the differential equation

$$\frac{dz}{ds} = f(s, z)$$

has a unique solution in D for $z : \mathbb{R} \to \mathbb{R}$ passing through

$$z(0) = \widehat{z}$$

and which extends to points arbitrarily close to the bundary of D.

(2) and

$$Y_t = nz(t/n) + o(n)$$

with probability at least $1 - n^{1/2}$ uniformly for $0 \leq t \leq \min\{\sigma n, m\}$, where $z(t)$ is the solution in (1) with $\widehat{z} = Y_0/n$, and $\sigma = \sigma(n)$ is the supremum of these s to which the solution can be extended.

Note that the bound in the probability in (2) is not directly stated in the original theorem of [Wor95], but was explicit in its proof.

Condition (i) ensures that the random variables Y_t does not change too drastically at each time step. Condition (ii) gives us an explicit target for this rate of change, while condition (iii) ensure that this rate of change does not evolve too quickly in time too.

Now suppose that we have a well-behaved supercritical sequence of degree sequence $\mathfrak{D} = \{\mathcal{D}_n\}_{n\geq 1}$. As per our exploration process, we expose a random configuration F on n vertices and degree sequence $\mathcal{D}_n$. The first step is to notice that, with high probability, the giant component is one of the first exposed.

Lemma 5.15. *For any function $\varphi(n) \to \infty$ such that $\varphi(n) = o(n/\log n)$, almost surely the largest component of F will be one of the first component exposed.*

Proof. For any configuration with degree sequence $\mathcal{D}_n$, we start by sorting the different components first by decreasing size of their edges sets, and then by decreasing highest labelled vertex. Let $\mathcal{C}$ be the smallest initial sequence of components with a total of at least ζn edges, where ζ is as defined in Theorem 5.4.

Let $\mathcal{E}$ be the event that one of the first $\varphi(n)$ exposed component is in $\mathcal{C}$. We start exploring a new component when all copies of partially exposed vertices have been used. Given that we have no vertex of

degree 1, the number of copies of the n^* unexposed vertices that are in $\mathcal{C}$ is at least $2\zeta n^*$, while on average we have a total of Kn^* copies of unexposed vertices (recall that K is a measure of average degree of $\mathcal{D}_n$). Therefore, the probability that a uniformly selected copy lies in $\mathcal{C}$ is at least $2\zeta/n$ and

$$\mathbb{P}[\mathcal{E}] \geq 1 - \left(1 - \frac{2\zeta}{K}\right)^{\varphi(n)}$$

This probability is equal to $1 - o(1)$ for $\varphi(n) = o(n/\log n)$.

Theorem 5.4 gives us that, in supercritical regime, the probability that $\mathcal{C}$ is composed of only the largest component is $1 - o(1)$. Therefore the probability that the largest component is one of the first $\varphi(n)$ exposed is also $1 - o(1)$, proving the lemma. □

The next corollary follows immediately,

Corollary 5.16. *Almost surely, the $\lfloor \log^2 n \rfloor^{th}$ edge edge exposed will be inside the largest component of F.*

We now follow on the proof of theorem 5.5.

Proof. We recall that during our exploration process, we defined $d_{i,j}$ to be the number of unexposed vertices of degree i after the exposition of j pairs. At start we have $d_{i,0} = n_i$ for all i. Fix $i_0 \geq 0$ and let $M = \sum_{i \geq 0} i n_i$ the initial total number of copies of vertices. When exposing the $(j+1)^{\text{th}}$ edge, the first vertex is chosen among open copies, hence has no impact on $d_{i,j}$, and then we can choose among $M - 2j - 1$ vertex copies in the set L, of which $i_0 d_{i_0,j}$ are copies of vertices with degree i_0. Therefore if $X_j > 0$ (i.e. there are some partially exposed vertices), the expected change in $d_{i_0,j}$ is

$$\mathbb{E}[d_{i_0,j+1} - d_{i_0,j}] = -\frac{i_0 d_{i_0,j}}{M - 2j - 1}$$

and the distribution of this change is mutually independent with the value of all $d_{i,j}$ with $i \neq i_0$.

Therefore, without the restriction $X_j > 0$, we could directly apply theorem 5.14 to $Y_j = d_{i_0,j}$ and, noting that $M/n = K + o(1)$, we could define the following function,

$$f(s, z) = \frac{-i_0 z}{K - 2s}$$

To remedy this complication, we :

(1) Start our analysis at step $j = \lfloor \log^2 n \rfloor$. Given corollary 5.16, almost surely, X_j will remain positive until the giant component has been fully exposed. However we still need to contain the case where X_j will drop to 0 prematurely.

(2) Introduce the variable $\delta_{i,j}$ to be the offset of $d_{i,j}$ if X_j stays positive along $\left[\lfloor \log^2 n \rfloor, j + \lfloor \log^2 n \rfloor\right]$, or if $j = 0$,

$$\delta_{i,j} = d_{i,j+\lfloor \log^2 n \rfloor}$$

otherwise

$$\delta_{i,j} = \begin{cases} \delta_{i,j-1} - 1 & \text{with probability } \frac{i\delta_{i,j-1}}{M - 2(j-1) - 1}, \\ \delta_{i,j-1} & \text{otherwise.} \end{cases}$$

Therefore, for a fixed $i_0 \geq 0$, using theorem 5.14, we can set $Y_j = \delta_{i_0,j}$. At each step j, the number of unexplored vertices varies by at most 1, hence we set $C' = 1$ and $m = n$. As seen above, if $X_j > 0$ stayed positive along $\left[\lfloor \log^2 n \rfloor, j + \lfloor \log^2 n \rfloor\right]$, then we can set $f(s,z) = \frac{-i_0 z}{K-2s}$. But if X_j drops earlier to 0, $|Y_{j+1} - Y_j| = 1$ with probability $\frac{i\delta_{i,j-1}}{M-2(j-1)-1}$ and $|Y_{j+1} - Y_j| = 0$ otherwise. Therefore, for any j,

$$\mathbb{E}[Y_{j+1} - Y_j \mid H_j] = f(j/n, Y_j/n) + o(1)$$

With $f(s,z) = \frac{-i_0 z}{K-2s}$. Applying theorem 5.14, we see that with probability at least $1 - n^{-1/2}$, for every $0 \leq \alpha = j/n < 1$,

$$\delta_{i_0,\lceil \alpha n \rceil} = nZ_{i_0}(\alpha) + o(n) \tag{29}$$

Where $Z_i(\cdot)$ is the unique solution to

$$\begin{cases} Z_i(0) = d_{i,0}/n \\ Z_i'(\alpha) = f(\alpha, Z_i(\alpha)) = -\frac{iZ_i(\alpha)}{K-2\alpha} \end{cases}$$

hence

$$Z_i(\alpha) = \frac{d_{i,0}}{n}\left(1 - \frac{2\alpha}{K}\right)^{i/2}$$

Note that $X_j = M - 2j - \sum_{i\geq 1} id_{i,j}$, after the exploration of j pairs the number of open copies is the total initial number of copies M, minus the $2j$ copies explored, minus the copies of unexplored vertices: $\sum_{i\geq 1} id_{i,j}$. Therefore,

$$X_{\lceil \alpha n \rceil} = M - 2\lceil \alpha n \rceil - \sum_{i\geq 1} id_{i,\lceil \alpha n \rceil}$$

Using corollary 5.16, almost surely,

$$X_{\lceil \alpha n \rceil} = Kn - 2\lceil \alpha n \rceil - \sum_{i\geq 1} id_{i,0}\left(1 - \frac{2\alpha}{K}\right)^{i/2} + S$$

Where $S = o(n)$, coming from $d_{i,\lfloor \log^2 n \rfloor} = d_{i,0} + o(n)$. Then, using the fact that $\mathfrak{D}$ is well-behaved (hence $d_{i,0} \approx \lambda_i n$),

$$X_{\lceil \alpha n \rceil} = \left(K - 2\alpha - \sum_{i\geq 1} i\lambda_i \left(1 - \frac{2\alpha}{K}\right)^{i/2}\right) n + S$$

Therefore, a.s,

$$X_{\lceil \alpha n \rceil} = \chi(\alpha)n + o(n)$$

This holds as long as $X_{\lceil \alpha n \rceil} > 0$, hence for all $0 \leq \alpha \leq \alpha_{\mathfrak{D}}$, and we have $X_{\lceil \alpha_{\mathfrak{D}} n \rceil} = o(n)$.

At this step, we stop our exploration process. In our configuration, the giant component has explored size $\sim 2\alpha_{\mathfrak{D}} n + o(n)$. By definition (see equation (25)) , α_D verifies:

$$2\alpha_{\mathfrak{D}} = K - \sum_{i\geq 1} i\lambda_i \left(1 - \frac{2\alpha_D}{K}\right)^{i/2}$$

Finally, we know from equation (29) that

$$\delta_{i,\lceil \alpha n \rceil} = d_{i,0}\left(1 - \frac{2\alpha}{K}\right)^{i/2} + o(n)$$

Almost surely, after $\lceil \alpha_{\mathfrak{D}} n \rceil$ steps, for each i, there are $\delta_{i,\lceil \alpha_{\mathfrak{D}} n \rceil}$ vertices of degree i unexplored. Therefore, using again the fact that $\mathfrak{D}$ is well-behaved (hence $d_{i,0} \approx \lambda_i n$), the giant component of G has explored order a.s. $\zeta_{\mathfrak{D}} n + o(n)$, with

$$\zeta_{\mathfrak{D}} n = n - \sum_{i \geq 1} \delta_{i,\lceil \alpha_{\mathfrak{D}} n \rceil}$$

$$\zeta_{\mathfrak{D}} = 1 - \sum_{i \geq 1} \lambda_i \left(1 - \frac{2\alpha_{\mathfrak{D}}}{K}\right)^{i/2}$$

Therefore, if we look at the graph obtained from G by removing the explored part, the methodology applied below for the discrete duality principle proves that we are left with a graph G' in subcritical regime, hence rapidly dying, proving that a.s. we are done exploring the giant component and it has order $\zeta_{\mathfrak{D}} n + o(n)$. □

5.6 Proof of Theorem 5.6 - Discrete duality principle

By lemma 5.15, for any $\omega(n) \to \infty$, we a.s. expose less than $\omega(n)$ components before exposing the giant component. In fact, with probability $\zeta_{\mathfrak{D}}$, the giant component is the first one exposed.

After completion of the exposure of the giant component, the configuration induced by the unexposed vertices is a uniformly random configuration with $d_{i,j}$ vertices of degree i, where j is the number of exposed pairs. By theorem 5.5, this configuration has a.s. $n' = (1 - \zeta_{\mathfrak{D}})n$ vertices.

Recall the definition of λ'_i (see equation (27)):

$$\lambda'_i = \frac{\lambda_i}{1 - \zeta_{\mathfrak{D}}}\left(1 - \frac{2\alpha}{K}\right)^{\frac{i}{2}}$$

Then

$$n(1 - \zeta_{\mathfrak{D}}) \sum_{i \geq 1} i\lambda'_i = n \sum_{i \geq 1} i\lambda_i \left(1 - \frac{2\alpha}{K}\right)^{\frac{i}{2}}$$

$$n' \sum_{i \geq 1} i\lambda'_i = n(K - 2\alpha_{\mathfrak{D}})$$

Therefore the new configuration verifies a.s. that for each i, $\lambda'_i n' + o(n')$ vertices have degree i.

Furthermore, note that as G has a.s. only one component of size greater than $\gamma \log n$ (by theorem 5.4), we should have $\mathcal{Q}(\mathfrak{D}) < 0$, which is easily verifiable :

For $0 < \alpha \leq \alpha_{\mathfrak{D}}$:

$$\begin{aligned}
\frac{d\chi(\alpha)}{d\alpha} &= -2 + (K-2\alpha)^{-1} \sum_{i\geq 1} i^2 \lambda_i \left(1 - \frac{2\alpha}{K}\right)^{i/2} \\
&= (K-2\alpha)^{-1} \left(\sum_{i\geq 1} i^2 \lambda_i \left(1 - \frac{2\alpha}{K}\right)^{i/2} - 2(K-2\alpha) \right) \\
&\geq (K-2\alpha)^{-1} \left(\sum_{i\geq 1} i^2 \lambda_i \left(1 - \frac{2\alpha}{K}\right)^{i/2} - 2\sum_{i\geq 1} i^2 \lambda_i \left(1 - \frac{2\alpha}{K}\right)^{i/2} \right) \\
&= (K-2\alpha)^{-1}(1-\zeta_{\mathfrak{D}}) \sum_{i\geq 1} i(i-2)\lambda'_i \\
&= (K-2\alpha)^{-1}(1-\zeta_{\mathfrak{D}})\mathcal{Q}(\mathfrak{D})
\end{aligned}$$

Hence, as the inequality is strict for $\alpha < \alpha_{\mathfrak{D}}$, and as α_D is the smallest positive zero of $\chi(\alpha)$, we must have $\mathcal{Q}(\mathfrak{D}) < 0$.

5.7 Other similar results

The results presented above are due to Molloy and Reed, and were the first of this type regarding the configuration model. These results have been strengthened in several ways. One may study the distribution of the error term in the result $L_1 = \zeta_{\mathfrak{D}} n + o(n)$; Among many others, see Kang and Seierstad [KS08], Pittel [Pit08], or Riordan [Rio12] for example. One may also ask for similar conclusion but with less restrictive assumptions :

(1) In [JL07], Janson and Luzczak replace the **MR**-conditions by the following ones (that we will call the **JL**-conditions).

JL1. $\mathfrak{D}$ is feasible smooth and sparse
JL2. $\sum_{i\geq 1} i^2 n_i = o(n)$
JL3. $\lambda = \sum_{i\geq 1} i\lambda_i$
JL4. $\lambda_1 > 0$

Note that if $\mathfrak{D}$ satisfies the **JL**-condition, by **JL**2, it is $O(n^{1/2})$-bounded. Janson and Luzczak showed a variant of Theorem 5.4 obtained with these conditions. They also extend the study to the case $\mathcal{Q}(\mathfrak{D}) = 0$.

(2) In [BR15], Bollobas and Riordan replace the **MR**-conditions by the following ones (that we will call the **BR**-conditions).

BR1. $\mathfrak{D}$ is feasible smooth and sparse
BR2. $\sum_{i\geq 3} \lambda_i > 0$
BR3. $\lambda = \sum_{i\geq 1} i\lambda_i$

Bollobas and Riordan showed a variant of Theorem 5.4 obtained with these conditions. They also extend the study to the case $\mathcal{Q}(\mathfrak{D}) = 0$. Their proof makes an extensive use of branching processes, proving some results on the distribution of the order of the giant component.

5.8 Return to Erdős-Rényi model

In [MR98], Molloy and Reed proved that the known results in the Erdős-Rényi model are special cases of Theorems 5.4 and 5.6.

The main idea is to transfer the random selection : for any graph $G \in G(n,p)$, start by exposing its degree sequence, and then choose a random graph with that degree sequence using the configuration model. Every graph with this degree sequence occurs with the same probability in $G(n,p)$, so this is a valid method of selection.

Then they use the following (well-known) lemma (seen in section 3.1),

Proposition. *For a random graph $G \in G(n,p)$ with $p = c/n$, a.s. for $i \leq O(\log n/\log\log n)$*

$$n_i = \frac{c^i}{i!}e^{-c}n + o(n),$$

and $n_i = 0$ otherwise.

Let $\zeta(\alpha)$ be defined* as

$$\begin{aligned}\zeta(\alpha) &= 1 - \sum_{i\geq 1}\frac{c^i}{i!}e^{-c}\left(1-\frac{2\alpha}{c}\right)^{i/2}\\ &= 1 - \exp(\sqrt{c^2-2c\alpha}-c)\end{aligned}$$

Then the size of the giant component given by (26) is $\zeta n + o(n)$ with $\zeta = \zeta(\alpha_{\mathfrak{D}})$, where $\alpha_{\mathfrak{D}}$ verifies equation (25),

$$c - 2\alpha_{\mathfrak{D}} - \sum_{i\geq 1} i\frac{c^i}{i!}e^{-c}\left(1-\frac{2\alpha_{\mathfrak{D}}}{c}\right)^{\frac{i}{2}}$$

And we can show that ζ is the unique solution to $\zeta + e^{-c\zeta} = 1$.

We can then prove that the Discrete Duality Principle (Theorem 3.11) seen in Erdős-Rényi model is a special case of Theorem 5.6. Let $d_c = c(1-\zeta)$. It is easy to show that $ce^{-c} = d_c e^{-d_c}$, hence

$$\frac{e^{-c}}{1-\zeta} = e^{-d_c}$$

We will show that if

$$\lambda_i = \frac{c^i}{i!}e^{-c},$$

Then

$$\lambda_i' = \frac{d_c^i}{i!}e^{-d_c}.$$

Since $\zeta = 1 - \exp(\sqrt{c^2-2c\alpha_{\mathfrak{D}}} - c)$, then $d_c = c\cdot\exp(\sqrt{c^2-2c\alpha_{\mathfrak{D}}} - c)$ and $c\exp(-c) = d_c\exp(-\sqrt{c^2-2c\alpha_{\mathfrak{D}}})$ and so

$$d_c = \sqrt{c^2-2c\alpha_{\mathfrak{D}}}$$

*we refer the reader to [MR98] for a full demonstration

Therefore

$$\begin{aligned}\lambda'_i &= \frac{\lambda_i}{1-\zeta_{\mathfrak{D}}}\left(1-\frac{2\alpha_{\mathfrak{D}}}{K}\right)^{\frac{i}{2}} \\ &= \frac{c^i e^{-c}}{i!(1-\zeta_{\mathfrak{D}})}\left(1-\frac{2\alpha_{\mathfrak{D}}}{K}\right)^{\frac{i}{2}} \\ &= \frac{e^{-c}}{i!(1-\zeta_{\mathfrak{D}})}\left(\sqrt{c^2-2\alpha_{\mathfrak{D}}c}\right)^i \\ &= \frac{d_c^i}{i!}e^{-d_c}.\end{aligned}$$

Hence verifying the discrete duality principle for the Erdős-Rényi model.

6. Switching

The studies on random graphs with given degree sequences presented above are all based on the configuration model. They provide results for some specific degree sequences (mainly smooth, sparse and well-behaved ones), under specific conditions (the so-called **MR**, **JL** or **BR**-conditions). These technical conditions do not allow the application of the presented results to many degree sequences that describe real-world networks such as for example *scale-free* networks*.

It is well-known† that many real-world networks are scale-free and one of the main research topic in this area is to determine the exponent of a particular network. It has been observed that many scale-free networks have a fat-tailed power-law degree distribution with exponent between 2 and 3. This is the case of the World Wide Web, where the exponent is between 2.15 and 2.2 ([FFF99]). In scale-free networks with exponents between 2 and 3, the vertices of high degree (called hubs) have a crucial role in several of the network properties such as in the *small-world* phenomenon. However, one of the main technical conditions under which the previous results on the existence of a giant component in $G(\mathcal{D})$ hold, is precisely that the vertices of high degree do not have a large impact on the structure of the graph. (In particular, they require the degree sequence to be sparse). Hence, often these results cannot be directly applied to real world networks where hubs are present and for each particular network ad-hoc approaches are needed (see for instance the Aiello-Chung-Lu model for the case of scale-free networks [ACL00]).

Unfortunately, without these technical conditions, during the exploration process the expected increase in X_i may change drastically, and checking that the initial value of $\mathcal{Q}(\mathfrak{D})$ is positive might not be sufficient to determine the existence of a giant component.

For instance let us consider the degree sequence $\mathcal{D}_n = (1, \dots, 1, d_n)$ with d_n being set either at $d_n = \lfloor 2\sqrt{n} \rfloor$ or at $\lfloor 2\sqrt{n} \rfloor + 1$ in order to ensure that $\sum_i d_i$ is even. Note that $\mathfrak{D} = \{\mathcal{D}_n\}_{n\geq 1}$ is feasible (by Erdős-Gallai theorem A.4), smooth (with $\lambda_1 = 1$, and $\lambda_i = 0$ for all $i \neq 1$) and sparse (again with $\lambda = 1$).

For simplicity, we consider the case $n = k^2$ with k odd so that $d_n = 2k$. The initial expected increase in X_i is

$$\frac{\mathcal{Q}(\mathfrak{D})}{K} = \frac{\sum_{k=1}^n d_k(d_k-2)}{\sum_{k=1}^n d_k} = \frac{4k^2-4k-(n-1)}{2k+n-1} \approx 3$$

*A network is scale-free if its degree distribution follows a power-law, governed by a specific exponent

†see [Hof16]

The approach from Molloy and Reed would suggest that $G(\mathcal{D}_n)$ a.a.s has a giant component. However, almost surely, the graph is composed of a star with $2k$ leaves, and $\frac{n-2k-1}{2}$ components K_2, hence without giant component (the largest component as size $\sim \sqrt{n} = o(n)$). This is because $\mathfrak{D}$:

- does not satisfy the **MR**2-conditions : it is not f-bounded by $f(n) = n^{1/4-\varepsilon}$.
- does not satisfy the **JL**2-conditions : $\sum_{i\geq 1} i^2 n_i \sim n \neq o(n)$.
- does not satisfy the **BR**2-conditions : $\sum_{i\geq 3} \lambda_i = 0$.

We cannot ensure that the expected increase in X_i stays positive during our exploration process : indeed as long as we only explore components K_2s, the expected increase is very stable (tending to 3), but as soon as we explore the n^{th}-vertex (with degree $d_n = 2k$), the expected increase of X_i drops to -1 until the star component is explored when the expected increase will be null.

In [JPRR18], Joos, Perarnau, Rautenbach and Reed extended the results of Molloy and Reed to arbitrary well-behaved degree sequences. Their work is based on a *switching* combinatorial argument.

Switching was introduced in the late 19th century by Petersen [Pet91] . Much later, McKay [McK11] reintroduced the method to count graphs with given degree sequences and, together with Wormald, used it in the study of random regular graphs in [Wor99].

The basic idea behind the method of switching is a double-counting argument: given two finite sets A, B and a relation R between them (the switching operation), the ratio of the average number of elements of B related to each element of A to the average number of elements of A related to each element of B is the same as the ratio of $|B|$ to $|A|$.

For any $a \in A$ and $b \in B$, if we denote by $R(a, b)$ the fact that the two element are related, then this is only using the fact that

$$\sum_{a\in A} \big|\{b \in B,\ R(a, b)\}\big| = \sum_{b\in B} \big|\{a \in A,\ R(a, b)\}\big|, \tag{30}$$

and therefore as stated,

$$\frac{\frac{1}{|A|}\sum_{a\in A} \big|\{b \in B,\ R(a, b)\}\big|}{\frac{1}{|B|}\sum_{b\in B} \big|\{a \in A,\ R(a, b)\}\big|} = \frac{|B|}{|A|}.$$

For each application the key problem is to find a suitable definition of R.

Now, let $\mathcal{A}$ and $\mathcal{B}$ be two disjoint sets of graphs. Suppose that for any graph $G \in \mathcal{A}$ there are $s_{\mathcal{A}\to\mathcal{B}}$ switchings transforming G into a graph in $\mathcal{B}$, while for any graph $G \in \mathcal{B}$ there are $s_{\mathcal{B}\to\mathcal{A}}$ switchings transforming G into a graph in $\mathcal{A}$. Then, using the double counting principle seen in equation (30) it follows immediately that,

$$|\mathcal{A}|s_{\mathcal{A}\to\mathcal{B}} = |\mathcal{B}|s_{\mathcal{B}\to\mathcal{A}}$$

Finally, if we can find an upper-bound $s^+ \geq s_{\mathcal{A}\to\mathcal{B}}$ and a lower-bound $s^- \leq s_{\mathcal{B}\to\mathcal{A}}$ we obtain that,

$$|\mathcal{A}|s^+ \geq |\mathcal{B}|s^- \tag{31}$$

This relation will be frequently used in the following work.

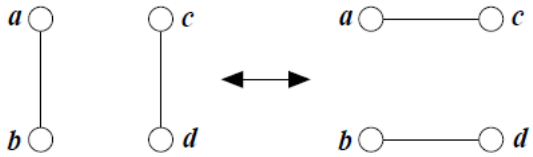


Figure 11: Switching example

6.1 Illustration

In order to grasp the potential of the switching methodology, we begin by proving that in any graph, if the number of edges is large with respect to the number of vertices, then there exists a component containing most of the vertices.

For a given degree sequence $\mathcal{D}$, the number of edges in $G(\mathcal{D})$ is given by $m = \frac{1}{2}\sum_{i=1}^{n} d_i$.

Proposition 6.1. *For any degree sequence $\mathcal{D}$, for any $\varepsilon > 0$, if $m = \omega(n)$, i.e if the number of edges dominates the number of vertices, then with high probability $G(\mathcal{D})$ has a component of order $(1-\varepsilon)n$.*

For this illustration, we will work on the simple switching seen in figure 11. Here, two graphs are related if there exists a switching transforming one into another.

We begin by proving the following lemma on 2-edge cuts of graphs,

Lemma 6.2. *Let G be any graph on n vertices and k components. Let S be the set of pairs of edges (uv, xy) such that by switching uv and xy we obtain a graph with $k+1$ components. Then $|S| = O(n^2)$.*

Proof. Let (uv, xy) be a pair of edges in S. Then clearly both edges must belong to the same connected component of G, otherwise the number of component of G could only decrease when applying the switching. As the function $x \mapsto x^2$ is convex, by simple Jensen inequality, if we upper bound $|S|$ in each component i of size n_i by $O(n_i^2)$, then the result follow for the whole graph G. Therefore we can suppose that G is connected.

First, suppose that at least one of the edges, w.l.o.g uv, is an edge cut of the graph. Then we claim that if xy is not an edge-cut too, the switching does not disconnect G. Indeed let uv be an edge cut of G into two connected components G_1 and G_2. W.l.o.g suppose that $u \in G_1$, and that $xy \in G_1$. For any $a, b \in G$ (see table 4 for visualization), denoting G' the graph obtained after switching (uv, xy) in G,

- if $a, b \in G_2$ then the path (ab) is not affected by our switching, and the path (ab) exists in G'.
- if $a, b \in G_1$ then because xy is not a cut-edge, there must exist a path in G_1 from a to b not using the edge xy. Therefore this path sill exists in G'.
- if $a \in G_1$ and $b \in G_2$. In G there is a path from b to v, hence in G' there is a path $(b \dots vy)$. But again because xy is not an edge cut, there must exist a path $(a \dots y)$ not going through x, still present in G'. Therefore in G' there is a path $(b \dots vy \dots a)$.

Therefore if uv is an edge cut of the graph then so must be xy in order to disconnect G.

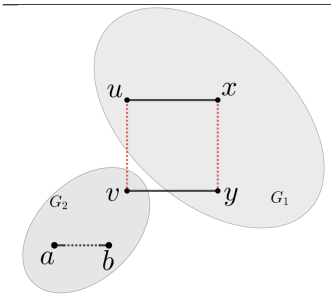
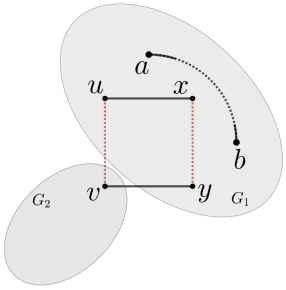
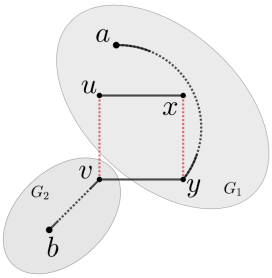
Before switching uv was a cut edge, but not xy. The switching does not disconnect G.		
		
$a, b \in G_2$	$a, b \in G_1$	$a \in G_1,\ b \in G_2$

Table 4: 3 possible cases for edge-cuts of lemma 6.2

Note that for any connected graph G on n vertices, there are at most $n-1$ cut-edges (the edges from any spanning tree of G). Therefore if uv is a cut-edge, there is at most $n-1$ edge xy in such that the pair (uv,xy) is in S. Therefore, knowing that there are at most $n-1$ cut-edge uv, we deduce that there are at most $(n-1)^2$ switchings using cut-edges.

Now, suppose that neither uv nor xy is a cut-edge, therefore (uv, xy) must be a proper 2-edge cut : Both $G-uv$ and $G-xy$ are connected, but $G-(uv,xy)$ is disconnected. Let T be a spanning tree of G. This tree must contain at least one edge from each 2-edge cut of G. Take uv among the $n-1$ edges of T. In order for (uv, xy) to be a 2-edge cut, we need xy to be an edge cut of $G-uv$, hence again $n-1$ choices possible and in total $(n-1)^2$ pairs of edges in S. Summing the two we obtain the desired bound. □

Using this lemma we can use the switching methodology to prove proposition 6.1.

Proof of proposition 6.1. We recall our statement : For any degree sequence $\mathcal{D}$, for any $\varepsilon > 0$, if $m = \omega(n)$ then with high probability $G(\mathcal{D})$ has a component of order $(1-\varepsilon)n$.

Our goal is to apply the methodology seen above in equation (31). To do so, fix $\varepsilon > 0$, and for any integer $k \geq 1$ let :

- $\mathcal{P}_k$ be the event that $G(\mathcal{D})$ has exactly k connected components,
- $\mathcal{Q}_k$, the event that $G(\mathcal{D})$ has exactly k connected components and each component has order at most $K = (1-\varepsilon)n$.

We denote $\mathcal{Q} = \bigcup_{k\geq 2} \mathcal{Q}_k$, and we want to prove that $\mathbb{P}[\mathcal{Q}] = o(1)$. This will imply that with high probability G has a component of order larger than K.

Fix some $k \geq 1$. We count the number of switchings between $\mathcal{P}_k$ and $\mathcal{Q}_{k+1}$. Any switching from $\mathcal{P}_k$ to $\mathcal{Q}_{k+1}$ must increase the number of connected components by 1. Therefore by lemma 6.2, they are at most $s^+ = O(n^2)$ such switchings.

Now let $G \in \mathcal{Q}_{k+1}$. Any switching occurring between a non-cut-edge uv and any another edge xy in a different connected component will decrease the number of connected component by 1. Using that the number of cut-edge in any graph is at most $n-1$, then the number of non-cut-edge uv is at least $m-(n-1)$. Then, since the component of uv has at most order $(1-\varepsilon)n$, then there are at least εn choices for xy. We obtain that we can lower bound the number of switching from $\mathcal{Q}_{k+1}$ to $\mathcal{P}_k$ by

$$s^- \geq (m-(n-1))\varepsilon n$$

And using $m = \omega(n)$ we obtain $s^- = \omega(n^2)$.

Then applying the switching principle :

$$\mathbb{P}[\mathcal{Q}_{k+1}]s^- \leq \mathbb{P}[\mathcal{P}_k]s^+$$

And

$$\mathbb{P}[\mathcal{Q}_{k+1}] \leq \frac{s^+}{s^-}\mathbb{P}[\mathcal{P}_k]$$
$$\mathbb{P}[\mathcal{Q}_{k+1}] \leq \frac{O(n^2)}{\omega(n^2)}\mathbb{P}[\mathcal{P}_k]$$
$$\mathbb{P}[\mathcal{Q}_{k+1}] \leq f(n)\mathbb{P}[\mathcal{F}_P],$$

with $f(n) = o(1)$.

Finally

$$\mathbb{P}[\mathcal{Q}] = \sum_{k\geq 1}\mathbb{P}[\mathcal{Q}_{k+1}] \leq f(n)\left(\sum_{k\geq 1}\mathbb{P}[\mathcal{P}_k]\right) = f(n) = o(1)$$

and with high probability, $G(\mathcal{D})$ has a component of order $(1-\varepsilon)n$. □

6.2 Presentation

We now present the specific switching that we will use in our giant component study.

For a given simple graph G, let us first consider the pair (G, H_G), where H_G is the multigraph obtained from G by deleting its cyclic components and suppressing the other vertices of degree 2. While the switching is actually performed on H_G, yielding another multigraph H', the process presented below allows for the construction of a simple graph G' such that $H' = H_{G'}$. We now describe for which switching in H_G we can obtain such an H' and how to do so.

Consider directed walks (either a path or a cycle) of G, corresponding to oriented edges in H_G, and let uv and xy be an ordered pair of oriented distinct edges in H_G. In order to yield a simple graph G', in G we select the associated ordered pair of directed walks (uv) and (xy) such that none of the following hold:

(1) there is an edge of G between u and x which forms neither uv nor xy, and the walk corresponding to uv has one edge,

(2) there is an edge of G between v and y which forms neither uv nor xy and the walk corresponding to xy has one edge,

(3) $u = x$ and the directed walk corresponding to uv has at most two edges, or

(4) $v = y$ and the directed walk corresponding to xy has at most two edges.

The switching then works as follow: From the edge uv in H_G, let $uw_1 \dots w_{r-1}v$ be the corresponding directed in G, and similarly let $xz_1 \dots z_{s-1}y$ be the directed walk corresponding to xy. Then, delete the edges $w_{r-1}v$ and xz_1 and add the edges $w_{r-1}x$ and vz_1 (see figure 12 for a simple example).

Remark that it is equivalent to switch the ordered pair (uv, xy) or the ordered pair (yx, vu), as we obtain the same graph G'. Therefore, given a pair of edges (uv, xy), we only consider the four following possible switchings : (uv, xy), (uv, yx), (vu, xy) and (vu, yx). They may also yield the same graph G' but we consider them all valid since it will be simpler to count them with these multiplicities.

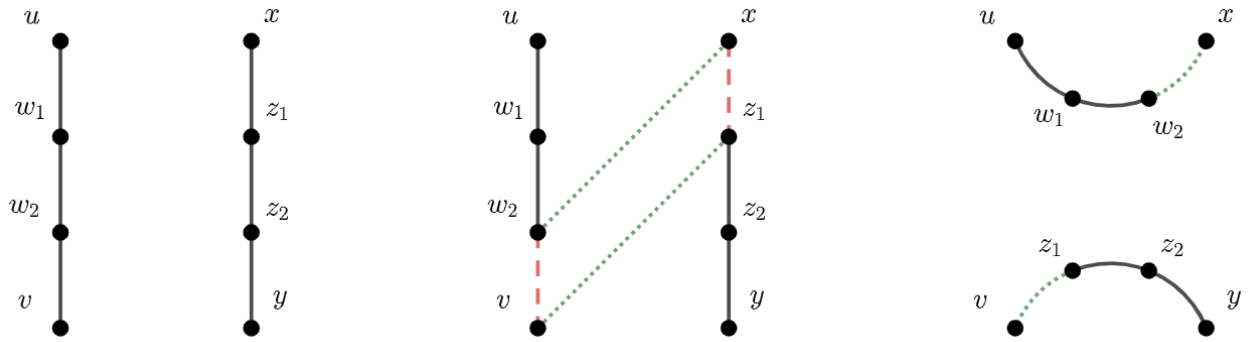


Figure 12: Switching for giant component study

6.3 Definitions and statement of results

Suppose that you are given a degree sequence $\mathcal{D}$. The original results from [JPRR18] used a permutation to order the degree sequence. In this presentation, for simplicity (mainly on notations) we will assume that $\mathcal{D} = (d_1, \dots, d_n)$ such that $d_1 \leq \dots \leq d_n$.

In order to generalize results presented above, let define the following invariant of $\mathcal{D}$:

- $j_{\mathcal{D}} = \min\left(\left\{j : j \in [n] \text{ and } \sum_{i=1}^{j} d_i(d_i - 2) > 0\right\} \cup \{n\}\right)$,
- $R_{\mathcal{D}} = \sum_{i=j_{\mathcal{D}}}^{n} d_i$,
- $M_{\mathcal{D}} = \sum_{\substack{i=1 \\ d_i \neq 2}}^{n} d_i$.

Intuitively, $j_{\mathcal{D}}$ represent a *breakpoint* between small and large degree vertices, note that $j_{\mathcal{D}} \geq 3$.Then $M_{\mathcal{D}}$ is the number of oriented edges in $H(\mathcal{D})$ and $R_{\mathcal{D}}$ is the number of edges attached to the large degree vertices. The expected increase of the number of open edges will remain positive until we have explored $R_{\mathcal{D}}$ edges, and it will then become negative. We will be able to explore a component with $R_{\mathcal{D}}$ edges.

For any function $\lambda : \mathbb{N} \to \mathbb{N}$, we say that a degree sequence $\mathcal{D}$ is well-behaved if $M_{\mathcal{D}}$ is at least $\lambda(n)$. The following results on degree sequences hold for any $\lambda \to \infty$ as $n \to \infty$.

Theorem 6.3. *For any function $\delta \to 0$ as $n \to \infty$, for every $\gamma > 0$, if $\mathcal{D}$ is a well behaved and graphic degree sequence with $R_{\mathcal{D}} < \delta(n) M_{\mathcal{D}}$, then the probability that $G(\mathcal{D})$ has a component of order at least γn is $o(1)$.*

Theorem 6.4. *For any positive constant ε, there is a $\gamma > 0$, such that if $\mathcal{D}$ is a well behaved and graphic degree sequence with $R_{\mathcal{D}} \geq \varepsilon M_{\mathcal{D}}$, then the probability that $G(\mathcal{D})$ has a component of order at least γn is $1 - o(1)$.*

Notwithstanding some technical complexities arising from vertices of degree 2*, the existence of a giant component will depend on whether or not $R_{\mathcal{D}}$ and $M_{\mathcal{D}}$ have same order. This additional complexity does

*We want to avoid situations where almost all vertices have degree 2

not impact the principle of our giant component analysis, but forces us to work first on the graph $H = H_G$ as detailed in the switching presentation: For a given simple graph G, H_G is the multigraph obtained from G by deleting its cyclic components and suppressing the other vertices of degree 2.

The proofs of theorems 6.3 and 6.4 both rely first on their equivalent for multigraphs $H(\mathcal{D})$:

Theorem 6.5. *For any function $\delta \to 0$ as $n \to \infty$, for every $\gamma > 0$, if $\mathcal{D}$ is a well behaved degree sequence with $R_{\mathcal{D}} < \delta(n)M_{\mathcal{D}}$, then the probability that $H(\mathcal{D})$ has a component of size at least $\gamma M_{\mathcal{D}}$ is $o(1)$.*

Theorem 6.6. *For any positive constant ε, there is a $\gamma > 0$, such that if $\mathcal{D}$ is a well behaved degree sequence with $R_{\mathcal{D}} \geq \varepsilon M_{\mathcal{D}}$, then the probability that $G(\mathcal{D})$ has a component of size at least $\gamma M_{\mathcal{D}}$ is $1 - o(1)$.*

In both phases (supercritical and subcritical), the proof of these theorems uses an exploration process combined with combinatorial and switching arguments. Then Joos & al. proved the following, allowing a transfer from the theorems on $H(\mathcal{D})$ to $G(\mathcal{D})$:

(1) For every $\gamma > 0$, there exists $\rho > 0$ such that for every well-behaved degree sequence $\mathcal{D}$, the probability that $G(\mathcal{D})$ has a component of order at least γn and $H(\mathcal{D})$ has no component of size at least $\rho M_{\mathcal{D}}$ is $o(1)$.

(2) For every $\rho > 0$, there exists $\gamma > 0$ such that for every well-behaved degree sequence $\mathcal{D}$, the probability that $G(\mathcal{D})$ has no component of order at least γn and $H(\mathcal{D})$ has a component of size at least $\rho M_{\mathcal{D}}$ is $o(1)$.

(3) For sequence that are not well-behaved, then the probability of having a giant component is bounded away both from 0 and 1.

We will give a complete proof of the subcritical case in section 6.5, and we refer to [JPRR18] in the supercritical case. As per the configuration model case, the aim in this later regime is to lower bound the expected increase in X_i, i.e the number of unexplored edges, and to prove that it is concentrated around its expected value.

6.4 Exploration process

Let $H(\mathcal{D})$ be the (multi)graph obtained from $G(\mathcal{D})$. In order to prove this concentration result, we need to ensure that at each step, the variation of our random variables is small. This usually means that we need to upper bound the degree of the vertices explored. We will do so by starting the exploration process with a non-empty set of vertices S_0, rather than a single vertex. With an appropriate choice for S_0, this allows us to handle vertices of high degree, avoiding encountering them during the exploration, and therefore allowing the expected concentration result.

In order to reduce the notation, we will often write G, H, V, M and R for $G(\mathcal{D})$, $H(\mathcal{D})$, $V(H(\mathcal{D}))$, $M_{\mathcal{D}}$ and $R_{\mathcal{D}}$.

At each step of our exploration process, we explore one vertex, so that it will take $|V \setminus S_0|$ steps to complete the exploration, defining

$$S_0 \subset S_1 \subset ... \subset S_{|V \setminus S_0|}$$

To help the exploration, for each vertex $v \in V$ we choose a uniformly random permutation of its adjacency list in G. The graph G together with the ordering of its adjacency lists for all $v \in V$ is called an **input**. Then, at each step t, we will define $w_t = S_t \setminus S_{t-1}$:

- If there is no edge between S_{t-1} and $V \setminus S_{t-1}$, then we select w_t at random in $V \setminus S_{t-1}$, with probability proportional to its degree.
- Otherwise,
  (1) we choose the smallest $v_t \in S_{t-1}$ (with respect to the natural order $\{1, \dots, n\}$) having a neighbour in $V \setminus S_{t-1}$.
  (2) With respect to the random permutation of its adjacency list, we explore the first edge between v_t and $V \setminus S_{t-1}$, and let w_t be the other endpoint.
  (3) Finally we expose all edges between w_t and S_{t-1}, as well as all loops of w_t.

As for the configuration exploration process, our main parameter will be X_t, the number of edges of H between S_t and $V \setminus S_t$, called **open edges**. Note that if $X_t = 0$, then S_t is the union of some fully explored components containing all S_0. Furthermore, if $|S_0| = 1$ then each X_i is a lower bound for the size of the largest component in H.

Since it is difficult to keep track of X_t, we define an auxiliary variable Z_t. We want Z_t to be an upper bound of X_t until $X_t = 0$. Note that at time $t = 0$, the number of open edges is at most the sum of degree of vertices in S_0, therefore we define

$$Z_0 = \sum_{u \in S_0} d(u) \geq X_0$$

Then, as long as $X_{t-1} > 0$, we will be able to find an edge $v_t w_t$ and we define an upper bound

$$Z_t = Z_0 + \sum_{i=1}^{t} (d(w_i) - 2) \geq X_t$$

Note that if S_0 is a stable set then $Z_0 = X_0$, and if furthermore the components containing the vertices of S_0 are all trees, then $Z_t = X_t$ as long as $X_t > 0$.

More importantly, if $Z_t = 0$, then there exists some $t' < t$ such that $X_{t'} = 0$, and the number of edges induced by S_t in H is a most $Z_t + 2t$.

6.5 Subcritical phase details

We recall that our goal is to prove Theorem 6.5,

Theorem. *For any function $\delta \to 0$ as $n \to \infty$, for every $\gamma > 0$, if $\mathcal{D}$ is a well behaved degree sequence with $R_{\mathcal{D}} < \delta(n) M_{\mathcal{D}}$, then the probability that $H(\mathcal{D})$ has a component of size at least $\gamma M_{\mathcal{D}}$ is $o(1)$.*

We will do so by proving that, under these conditions, for any vertex $v \in V$, we can start our exploration at some $S_0(v)$ containing v, such that the probability that there is some t with $X_t = 0$ and with at most γM edges in S_t is $1 - o(M^{-1})$. Therefore since H has at most $2M$ edges, then the probability to have a component of size at least γM tends to zero.

We start by stating our overall result, from which theorem 6.5 follows:

Lemma 6.7. *For every sufficiently small $\omega > 0$ and every degree sequence $\mathcal{D}$ such that $R < \omega M$ and M is sufficiently large in terms of ω, for every vertex $v \in H$, the probability that v lies in a component of H of size larger than $\omega^{1/9} M$ is less than $e^{-M^{1/4}}$.*

←-- $M_{\mathcal{D}} = \sum d_i$ --→														
d_1	$\leq$	...	$\leq$	$d_{j^*_{\mathcal{D}}}$	$\leq$	...	$\leq$	$d_{j_{\mathcal{D}}-1}$	$\leq$	$d_{j_{\mathcal{D}}}$	$\leq$	...	$\leq$	d_n
←-- Def $j_{\mathcal{D}}$: $\sum d_i(d_i - 2) \leq 0$ --→														
d_1	$\leq$	...	$\leq$	$d_{j^*_{\mathcal{D}}}$	$\leq$	...	$\leq$	$d_{j_{\mathcal{D}}-1}$	$\leq$	$d_{j_{\mathcal{D}}}$	$\leq$	...	$\leq$	d_n
←-- Def $j_{\mathcal{D}}$: $\sum d_i(d_i - 2) > 0$ --→														
										←-- $R_{\mathcal{D}} = \sum d_i < \omega M_{\mathcal{D}}$ --→				
d_1	$\leq$	...	$\leq$	$d_{j^*_{\mathcal{D}}}$	$\leq$	...	$\leq$	$d_{j_{\mathcal{D}}-1}$	$\leq$	$d_{j_{\mathcal{D}}}$	$\leq$	...	$\leq$	d_n
				←-- Def set S and $j^*_{\mathcal{D}}$: $\sum d_i \geq 5\omega^{1/4} M$ --→										
d_1	$\leq$	...	$\leq$	$d_{j^*_{\mathcal{D}}}$	$\leq$	...	$\leq$	$d_{j_{\mathcal{D}}-1}$	$\leq$	$d_{j_{\mathcal{D}}}$	$\leq$	...	$\leq$	d_n
						←-- Def set S and $j^*_{\mathcal{D}}$: $\sum d_i < 5\omega^{1/4} M$ --→								

Table 5: Definitions, in ordered $\mathcal{D}$, with $3 \leq d_{j_{\mathcal{D}}}$

We know that the number of edges induced by S_t in H is a most $Z_t + 2t$. Therefore it is enough to show that the probability that there is no t with $Z_t = 0$ and $t \leq \frac{\omega^{1/9} M}{2}$ is less than $e^{-M^{1/4}}$.

In this section, we implicitly assume that ω is small enough, and M large enough with respect to ω ($\omega M \gg 1$) to ensure that various inequalities (such that $c\omega^{1/4} < \omega^{1/5}$) scattered in the proofs are all satisfied.

We first give a rationale for our definition of S_0, giving an outline of the proof: As long as $X_t > 0$, the probability that a specific vertex w is chosen as w_t is essentially proportional to its degree. Therefore, the expected value of $Z_t - Z_{t-1} = d(w_t) - 2$ is with high probability,

$$\frac{\sum_{w \in V \setminus S_{t-1}} d(w)(d(w) - 2)}{\sum_{w \in V \setminus S_{t-1}} d(w)}$$

Therefore if we set S_0 as containing v and all high degree vertices (with respect to some breakpoint notion defined below, in line with our definition of $j_{\mathcal{D}}$), we can ensure that the initial expected increase $Z_1 - Z_0$ is negative. Then we will be able to bound the expected increase $Z_t - Z_{t-1}$ which decreases as t increases, allowing us to find the desired result.

Let S be the smallest set of vertices such that $\sum_{i \in S} d_i \geq 5\omega^{1/4} M$, and there is no vertex outside of S with degree higher than any vertex of S. Then we can define* $j^*_{\mathcal{D}}$ such that

$$\sum_{i \in S} d_i = \sum_{i = j^*_{\mathcal{D}}}^{n} d_i$$

Using the hypothesis of lemma 6.7, $R < \omega M$, and for sufficiently small ω, $\omega M < 5\omega^{1/4} M$, we obtain that $j^*_{\mathcal{D}} < j_{\mathcal{D}}$, and therefore by definition of $j^*_{\mathcal{D}}$

$$\sum_{w \in V \setminus S} d(w)(d(w) - 2) \leq 0$$

*We refer to table 5 for a quick visualisation of our definitions

However we can obtain a straightening of this inequality as follows. First, as explained, we let $S_0 = S \cup \{v\}$, and by definition of Z_0,

$$Z_0 = \sum_{u \in S_0} d(u) = \begin{cases} \sum_{u \in S} d(u) & \text{if } v \in S \\ d(v) + \sum_{u \in S} d(u) & \text{if } v \notin S \end{cases}$$

Note that because S is the smallest set of vertices such that $\sum_{i \in S} d_i \geq 5\omega^{1/4}M$, we must have that the sum of degree in S is at most $5\omega^{1/4}M + \delta(S)$ where $\delta(S)$ is the minimum degree in S.

$$Z_0 \leq \begin{cases} 5\omega^{1/4}M + \delta(S) & \text{if } v \in S \\ d(v) + 5\omega^{1/4}M + \delta(S) & \text{if } v \notin S \end{cases}$$

Now, suppose that in S there is a vertex of degree 1 (i.e. $d_{j^*_{\mathcal{D}}} = 1$). Then $\delta(S) = 1$ and any vertex in $V \setminus S$ has degree 1, hence

$$Z_0 \leq 5\omega^{1/4}M + 2$$

Therefore $Z_0 \leq \frac{\omega^{1/9}M}{2}$. But if every vertex in $V \setminus S$ has degree 1, then every edge in the component of H containing S_0 is incident to a vertex of S_0, hence there are at most Z_0 edges between S_0 and $V \setminus S_0$. Finally this implies that there are at most $2Z_0 \leq \omega^{1/9}M$ edges in the components containing S_0, and lemma 6.7 is proven.

Therefore from now on, suppose that every vertex in S has degree at least 3. Using $j^*_{\mathcal{D}} < j_{\mathcal{D}}$,

$$\begin{aligned} \sum_{w \in V \setminus S} d(w)(d(w)-2) &= \sum_{i=1}^{j^*_{\mathcal{D}}-1} d_i(d_i-2) \\ &= \underbrace{\sum_{i=1}^{j_{\mathcal{D}}-1} d_i(d_i-2)}_{<0 \text{ by minimality of } j_{\mathcal{D}}} - \sum_{i=j^*_{\mathcal{D}}}^{j_{\mathcal{D}}-1} d_i(d_i-2) \\ &\leq -\sum_{i=j^*_{\mathcal{D}}}^{j_{\mathcal{D}}-1} d_i(d_i-2) \end{aligned}$$

Now for all $i \in \{j^*_{\mathcal{D}}, \dots, j_{\mathcal{D}}\}$, we have $i \in S$, and therefore $d_i \geq 3$. Then

$$\begin{aligned} \sum_{w \in V \setminus S} d(w)(d(w)-2) &\leq -\sum_{i=j^*_{\mathcal{D}}}^{j_{\mathcal{D}}-1} d_i(d_i-2) \\ &\leq -\sum_{i=j^*_{\mathcal{D}}}^{j_{\mathcal{D}}-1} d_i = -\left(\sum_{i \in S} d_i - R\right) \\ &\leq -(5\omega^{1/4}M - R) \\ &\leq -4\omega^{1/4}M \end{aligned} \tag{32}$$

Furthermore note that for sufficiently small ω we have $\omega^{-1/4} - 2 > \frac{3\omega^{-1/4}}{4}$. Therefore if all vertices in S have degree at least $\omega^{-1/4}$, then we have

$$\sum_{i=j^*_{\mathcal{D}}}^{j_{\mathcal{D}}-1} d_i(d_i-2) > \frac{3\omega^{-1/4}}{4} \sum_{i=j^*_{\mathcal{D}}}^{j_{\mathcal{D}}-1} d_i$$

Then, using as we have seen in deduction (32) that $\sum_{i=j_{\mathcal{D}}^*}^{j_{\mathcal{D}}-1} d_i \geq 4\omega^{1/4}M$, we deduce that

$$\sum_{i=j_{\mathcal{D}}^*}^{j_{\mathcal{D}}-1} d_i(d_i-2) > \frac{3\omega^{-1/4}}{4} 4\omega^{1/4}M = 3M$$

Finally we obtain

$$\sum_{i=1}^{j_{\mathcal{D}}-1} d_i(d_i-2) > 3M + \underbrace{\sum_{i=1}^{j_{\mathcal{D}}^*-1} \underbrace{d_i(d_i-2)}_{>-1}}_{>-M}$$

and

$$\sum_{i=1}^{j_{\mathcal{D}}-1} d_i(d_i-2) > 2M > 0$$

This is a contradiction with the minimality of $j_{\mathcal{D}}$. In S there must be a vertex with degree at most $\omega^{1/4}$. As S gather the vertices of high degrees, this implies that

$$\forall u \in V \setminus S_0,\ d(u) \leq \omega^{-1/4}. \tag{33}$$

Then because S is the smallest set of vertices such that $\sum_{i\in S} d_i \geq 5\omega^{1/4}M$, as we have seen before in the case where S had a vertex of degree 1, we must have that the sum of degree in S is at most $5\omega^{1/4}M + \delta(S)$. Then

$$\begin{aligned} Z_0 &= \sum_{u\in S_0} d(u) \\ Z_0 &\leq d(v) + 5\omega^{1/4}M + \delta(S) \\ Z_0 &\leq 5\omega^{1/4}M + 2\omega^{-1/4} \\ Z_0 &\leq 7\omega^{1/4}M \end{aligned} \tag{34}$$

In our exploration process, we define the auxiliary variable :

$$Y_t = Z_t - Z_{t-1} - \mathbb{E}[d(w_t) - 2] = d(w_t) - \mathbb{E}[d(w_t)]$$

The following lemma reduce the problem of bounding Z_t to bounding $\mathbb{E}[d(w_t)]$:

Lemma 6.8. *The probability that there is a t such that $\sum_{t'\leq t} Y_{t'} > M^{2/3}$ is less than $e^{-M^{1/4}}$*

Proof. Recall Azuma's inequality (see A.5).

Proposition (Azuma's inequality)**.** *Let X be a random variable determined by a sequence of N random experiments $T_1, \dots, T_n$ such that for every $1 \leq i \leq N$ and for any possible sequences $t_1, \dots, t_{i-1}, t_i$ and $t_1, \dots, t_{i-1}, t_i'$ such that*

$$\left|\mathbb{E}[X \mid T_1 = t_1, \dots, T_i = t_i] - \mathbb{E}[X \mid T_1 = t_1, \dots, T_i = t_i']\right| \leq c_i$$

for some $c_i > 0$. Then

$$\mathbb{P}\big[|X - \mathbb{E}[X]| > t\big] < 2e^{-\frac{t^2}{2\sum_{i=1}^N c_i^2}}$$

We recall equation (33), stating that for any vertex u out of S_0, we have an upper bound for its degree: $d(u) \leq \omega^{-1/4}$. Therefore during our exploration process, for any t, $|Y_t| \leq \omega^{-1/4}$. We apply Azuma's inequality to the random variable $\sum_{t' \leq t} Y_{t'}$, with $\mathbb{E}[Y_t] = 0$, $N = t$ and $c_i = \omega^{-1/4}$, yielding

$$\mathbb{P}\left[\sum_{t' \leq t} Y_{t'} > M^{2/3}\right] < 2e^{-\frac{M^{4/3}}{2t\omega^{-1/2}}} < e^{-M^{2/7}},$$

Using the union bound over all $t \leq M$ we obtain the desired result. □

We let $M_t = \sum_{v \in V \setminus S_t} d(v)$. Now that we have reduced the problem of bounding Z_t to bounding $\mathbb{E}[d(w_t)]$, the main lemma to prove Theorem 6.5 is the following,

Lemma 6.9. *It $t \leq \omega^{1/9} M$, $Z_{t-1} \leq \omega^{1/5} M$, and $X_{t'} > 0$ for all $t' < t$, then:*

a. If $w \in V \setminus S_{t-1}$ and $d(w) = 1$ then

$$\mathbb{P}[w_t = w] \geq \left(1 - 9\omega^{1/5}\right) \frac{1}{M_{t-1}}.$$

b. If $w \in V \setminus S_{t-1}$ then

$$\mathbb{P}[w_t = w] \leq \left(1 + 9\omega^{1/5}\right) \frac{d(w)}{M_{t-1}}.$$

As the proof of this lemma uses our switching methodology, we defer it (as well as the proof of the next lemma) to a dedicated section.

Applying iteratively this result, we will bound the expected degree of w_t,

Lemma 6.10. *Let ρ be defined by*

$$\rho = \arg\min_t \left\{ \sum_{t' < t} Y_{t'} > M^{2/3} \text{ or } X_t = 0 \right\}$$

Then, with $\tau = \min\left\{\rho, \left\lfloor \frac{\omega^{1/9} M}{2} \right\rfloor\right\}$, for any $t \leq \tau$ we have:

$$\mathbb{E}[d(w_t) - 2] \leq -\frac{t}{M} + 19\omega^{1/5}$$

Finally the following lemma complete the proof of lemma 6.7, and therefore of Theorem 6.5.

Lemma 6.11. *With probability greater than $1 - e^{-M^{1/4}}$, there exist some $t \leq \left\lceil \frac{\omega^{1/9} M}{3} \right\rceil$ such that $X_t = 0$.*

Proof. Recall lemma 6.8 : The probability that there is a t such that $\sum_{t' \leq t} Y_{t'} > M^{2/3}$ is less than $e^{-M^{1/4}}$. Therefore with probability at least $1 - e^{-M^{1/4}}$, there is no t such that $\sum_{t' \leq t} Y_{t'} > M^{2/3}$.

Suppose additionally that our lemma does not hold : there is no $t \leq \left\lceil \frac{\omega^{1/9}M}{3} \right\rceil$ such that $X_t = 0$. Then, when applying lemma 6.10, we obtain $\tau \geq \left\lceil \frac{\omega^{1/9}M}{3} \right\rceil$, and that, for all $t \leq \left\lceil \frac{\omega^{1/9}M}{3} \right\rceil$,

$$\begin{aligned} X_t \leq Z_t &= Z_0 + \sum_{i=1}^{t} (d(w_i) - 2) \\ &\leq Z_0 + \sum_{i=1}^{t} (Y_i + \mathbb{E}[d(w_i) - 2]) \\ &\leq Z_0 + M^{2/3} + \sum_{i=1}^{t} \left(-\frac{i}{M} + 19\omega^{1/5} \right) \\ &\leq Z_0 + M^{2/3} - \frac{t(t-1)}{2M} + 19\omega^{1/5} t \end{aligned}$$

Finally using equation (34),

$$X_t \leq 7\omega^{1/4} + M^{2/3} - \frac{t(t-1)}{2M} + 19\omega^{1/5} t$$

And for $t = \left\lceil \frac{\omega^{1/9}M}{3} \right\rceil$ we obtain $X_t < 0$, a contradiction. □

6.5.1 Proof of lemma 6.9

First we lower bound the number of vertices of degree 1.

$$\begin{aligned} \sum_{i=1}^{j_{\mathcal{D}}-1} d_i(d_i - 2) &= -n_1 + \sum_{\substack{i=1 \\ d_i \neq 1}}^{j_{\mathcal{D}}-1} d_i(d_i - 2) \\ &\geq -n_1 + \sum_{\substack{i=1 \\ d_i \neq 1 \\ d_i \neq 2}}^{j_{\mathcal{D}}-1} d_i \\ &\geq -2n_1 + \sum_{\substack{i=1 \\ d_i \neq 2}}^{j_{\mathcal{D}}-1} d_i \end{aligned}$$

Hence by definition of $j_{\mathcal{D}}$:

$$0 \geq -2n_1 + \sum_{\substack{i=1 \\ d_i \neq 2}}^{j_{\mathcal{D}}-1} d_i$$

Hence, $n_1 \geq \frac{M-R}{2}$ and by hypothesis of lemma 6.9, $n_1 \geq (1-\omega)M \geq \frac{M}{3} + 1$. Now since $t < \frac{M}{12}$ (for small enough ω as $t < \omega^{1/9}M$) then we know that the set $V \setminus S_{t-1}$ contains at least $\frac{M}{4} + 1$ of the $\frac{M}{3} + 1$ vertices of degree 1,

$$M_{t-1} \geq \frac{M}{4} + 1 \tag{35}$$

Note that if there is no edges between S_{t-1} and $V \setminus S_{t-1}$, then

$$\mathbb{P}[w_t = w] = \frac{d(w)}{M_{t-1}}$$

Hence the two parts of our lemma hold.

Now assume that there is at least an edge between S_{t-1} and $V \setminus S_{t-1}$. We define the following equivalence class on all inputs* : Given two inputs I_1, I_2,

$$I_1 \sim I_2 \text{ if and only if they share } \begin{cases} \text{the same underlying graph } G \\ \text{the same ordering of its adjacency list for each vertex in } S_{t-1} \\ \text{the same edge } v_t w_t \end{cases}$$

And we'll note

$$[I_1]_\sim = \{I \mid I_1 \sim I\}$$

We define the two disjoints sets of equivalence classes :

$$\mathcal{A}_w = \left\{[I]_\sim,\ w = w_t\right\}$$

$$\mathcal{B}_w = \left\{[I]_\sim,\ w \neq w_t\right\}$$

We want to perform switching between these two sets. To do so we need to ensure that our switching preserve the equivalence classes (i.e. if two inputs are equivalent, they still are after our switching operation). To ensure this we restrict the switching to the ones that are neither using nor creating any edge in S_t.

Proof of 6.9a. Suppose that $d(w) = 1$. Any switching from $\mathcal{A}_w$ to $\mathcal{B}_w$ must involve

- either the edge $v_t w$ and one of the remaining $M_{t-1} - 1$ oriented edges xy with $x \in V \setminus S_{t-1}$
- or the edge wv_t and one of the remaining $M_{t-1} - 1$ oriented edges xy with $y \in V \setminus S_{t-1}$

Therefore there are at most $4M_{t-1}|\mathcal{A}_w|$ switching from $\mathcal{A}_w$ to $\mathcal{B}_w$.

We need now to lower bound the number of switching from $\mathcal{B}_w$ to $\mathcal{A}_w$. Each choice of equivalence class in $\mathcal{B}_w$ define the edge $v_t w_t$. We will call any vertex in $V \setminus S_{t-1}$ *bad* if

- it has a neighbour in S_{t-1},
- or has a common neighbour with w_t,
- or is a neighbour of w_t,

Otherwise, it is *good*, and with z its unique neighbour, there are 4 switching from $\mathcal{B}$ to $\mathcal{A}_w$ using the pair $(v_t w_t, wz)$. By hypothesis of our lemma, we have that $Z_t < \omega^{1/5} M$ and as Z_t upper-bounds X_t we know that there are at most $\omega^{1/5} M$ vertices in $V \setminus S_{t-1}$ with a neighbour in S_{t-1}. Additionally, by equation (33), we know that vertices in $V \setminus S_{t-1}$ have degree at most $\omega^{-1/4}$, therefore there are at most $\omega^{1/5} M + 2\omega^{-1/4} \leq 2\omega^{1/5} M$ bad vertices (again for ω small enough). As we have at least $M/4$ vertices

*We recall that an input is the graph G together with the ordering of its adjacency lists for all $v \in V$

of degree 1 in $V \setminus (S_{t-1} \cup \{v_t\})$, then the proportion of equivalence classes in $\mathcal{B}_w$ such that w is a good vertices is at least $1 - \frac{2\omega^{1/5}M}{M/4} = 1 - 8\omega^{1/5}$. Hence there are at least $4(1 - 8\omega^{1/5})|\mathcal{B}_w|$ switching from $\mathcal{B}_w$ to $\mathcal{A}_w$.

The double counting methodology returns :

$$|\mathcal{A}_w| \geq \frac{(1 - 8\omega^{1/5})|\mathcal{B}_w|}{M_{t-1}}$$

Finally, using that $\mathbb{P}[w_t = w] = \frac{|\mathcal{A}_w|}{|\mathcal{A}_w|+|\mathcal{B}_w|}$, and $M_{t-1} = M - Z_{t-1} - 2(t-1)$, it follows that

$$\mathbb{P}[w_t = w] \geq \frac{1 - 9\omega^{1/5}}{M_{t-1}}$$

□

Proof of 6.9b. Let $d = d(w)$. For a switching from $\mathcal{B}_w$ to $\mathcal{A}_w$ we need to switch the ordered edge $v_t w_t$ with one of the d ordered edges wy or the ordered edge $w_t v_t$ with one of the d ordered edges xw, therefore there are at most $4d|\mathcal{B}_w|$ switching from $\mathcal{B}_w$ to $\mathcal{A}_w$.

Now suppose that $v_t y$ and wz are not edges. The switching between the ordered edge $v_t w$ and some ordered edge yz with y in $V \setminus S_{t-1}$ is a switching from $\mathcal{A}_w$ to $\mathcal{B}_w$. And so is the switching between the ordered edge wv_t and some ordered edge zy with y in $V \setminus S_{t-1}$. Since v_t has degree at most ωM, and the maximum degree of any vertex in $V \setminus S_{t-1}$ is $\omega^{-1/4}$, then the number of choices for zy is at least $M_{t-1} - \omega^{3/4}M - Z_{t-1} - \omega^{-1/2} > (1 - 2\omega^{1/5})\frac{M_{t-1}}{d}|\mathcal{A}_w|$. Using again that $\mathbb{P}[w_t = w] = \frac{|\mathcal{A}_w|}{|\mathcal{A}_w|+|\mathcal{B}_w|}$ we obtain

$$\mathbb{P}[w_t = w] \leq \frac{1 + 9\omega^{1/5}}{M_{t-1}}$$

□

6.5.2 Proof of lemma 6.10

We prove our statement by induction on t.

Base case $t = 1$**.** Recall that by equation(34),

$$Z_0 \leq 7\omega^{1/4}M \leq \frac{\omega^{1/5}M}{2}$$

Therefore we can apply lemma 6.9, and can compute the expected degree of w_1 :

$$\begin{aligned}
\mathbb{E}[d(w_1)-1] &= \sum_{w\in V\setminus S_0}(d(w)-2)\mathbb{P}[w_1=w] \\
&= \sum_{\substack{w\in V\setminus S_0\\ d(w)\geq 3}}(d(w)-2)\mathbb{P}[w_1=w] - \sum_{\substack{w\in V\setminus S_0\\ d(w)=1}}\mathbb{P}[w_1=w] \\
&\leq (1+9\omega^{1/5})\frac{1}{M_0}\sum_{\substack{w\in V\setminus S_0\\ d(w)\geq 3}}(d(w)-2)d(w) - (1-9\omega^{1/5})\frac{1}{M_0}\sum_{\substack{w\in V\setminus S_0\\ d(w)=1}}1 \\
&\leq (1+9\omega^{1/5})\frac{1}{M_0}\sum_{\substack{w\in V\setminus S_0\\ d(w)\geq 3}}(d(w)-2)d(w) + (1-9\omega^{1/5})\frac{1}{M_0}\sum_{\substack{w\in V\setminus S_0\\ d(w)=1}}(d(w)-2) \\
&\leq (1+9\omega^{1/5})\frac{1}{M_0}\sum_{w\in V\setminus S_0}(d(w)-2)d(w) - 18\omega^{1/5}\frac{1}{M_0}\sum_{\substack{w\in V\setminus S_0\\ d(w)=1}}(d(w)-2) \\
&\leq (1+9\omega^{1/5})\frac{1}{M_0}\sum_{w\in V\setminus S_0}(d(w)-2)d(w) + 18\omega^{1/5}\frac{1}{M_0}n_1 \\
&\leq \frac{1}{M_0}\left(18\omega^{1/5}n_1 + (1+9\omega^{1/5})\sum_{w\in V\setminus S_0}(d(w)-2)d(w)\right)
\end{aligned}$$

Now recall that by equation 32

$$\sum_{w\in V\setminus S}d(w)(d(w)-2) \leq -4\omega^{1/4}M$$

Therefore

$$\mathbb{E}[d(w_1)-1] \leq \frac{1}{M_0}\left(18\omega^{1/5}n_1 - (1+9\omega^{1/5})\cdot 4\omega^{1/4}M\right)$$

Now, as there is no vertex of degree 1 in S, then there is at most 1 vertex of degree 1 in S_0 (the initial vertex v), hence $n_1 \leq M_0+1$ and

$$\mathbb{E}[d(w_1)-1] \leq (18\omega^{1/5})\frac{M_0+1}{M_0} - \frac{(1+9\omega^{1/5})\cdot 4\omega^{1/4}M}{M_0}$$

Finally, using that $M_0 \leq M$ and that $\omega M \gg 1$, we obtain

$$\mathbb{E}[d(w_1)-1] \leq 19\omega^{1/5} - \frac{1}{M}$$

Induction $2 \leq t \leq \frac{\omega^{1/9}M}{2}$.

By induction on $\sum a_i$, one can easily prove* the following :

Lemma 6.12. *For any sequence of positive integers $a_1, \dots, a_j$ distinct from 2 and a non-negative integer ℓ, such that $\sum_{i=1}^{j} a_i \geq 2j-\ell$, we have $\sum_{i=1}^{j} a_i(a_i-2) \geq j-2\ell$*

*see claim 1 in [JPRR18]

Since $X_{t-1} > 0$, we know that $Z_{t-1} > 0$, and hence

$$\sum_{i=1}^{t-1} d(w_i) = 2(t-1) + \sum_{i=1}^{t-1}(d(w_i)-2) = 2(t-1) + Z_{t-1} - Z_0 \geq 2(t-1) - Z_0$$

Therefore applying this lemma we obtain that

$$\sum_{i=1}^{t-1} d(w_i)(d(w_i)-2) \geq (t-1) - 2Z_0$$

Since $V \setminus S_{t-1} = V \setminus (S_0 \cup \{w_1, \dots, w_{t-1}\})$, with χ_v the characteristic function of $v \notin S$, we know that

$$\begin{aligned}\sum_{w \in V \setminus S_{t-1}} d(w)(d(w)-2) &= \sum_{w \in V \setminus S_0} d(w)(d(w)-2) - \sum_{i=1}^{t-1} d(w)(d(w)-2) \\ &\leq \sum_{w \in V \setminus S} d(w)(d(w)-2) - \underbrace{d(v)(d(v)-2)\chi_v}_{\geq -1} + 2Z_0 - (t-1)\end{aligned}$$

Then by equation (32), using again $\omega M \gg 1$,

$$\begin{aligned}\sum_{w \in V \setminus S_{t-1}} d(w)(d(w)-2) &\leq -4\omega^{1/4}M + 1 + 2Z_0 - (t-1) \\ \sum_{w \in V \setminus S_{t-1}} d(w)(d(w)-2) &\leq 2Z_0 - (t-1) \qquad (36)\end{aligned}$$

Now, by definition we have

$$Z_{t-1} = Z_0 + \sum_{i<t}(d(w_i)-2)$$

Using $Y_t = d(w_t) - \mathbb{E}[d(w_t)]$ we deduce that

$$Z_{t-1} = Z_0 + \sum_{i<t}(\mathbb{E}[d(w_i)]-2) + \sum_{i<t} Y_i$$

By induction principle, and using the lemma's hypothesis $\sum_{i<t} Y_i \leq M^{2/3}$,

$$Z_{t-1} = Z_0 + \sum_{i<t}\left(-\frac{i}{M} + 19\omega^{1/5}\right) + M^{2/3}$$

Finally, with $t \leq \frac{\omega^{1/9}M}{2}$ and using $Z_0 \leq \frac{\omega^{1/5}M}{2}$ we obtain that

$$Z_{t-1} \leq \omega^{1/5}M$$

Therefore we can apply lemma 6.9. Combining it with equation (36), as per the base case, we obtain

$$\mathbb{E}[d(w_t)-2] \leq \frac{1}{M_{t-1}}\left((1+9\omega^{1/5})(2Z_0-(t-1)) + 18\omega^{1/5}n_1\right)$$

And since $n_1 \leq M_{t-1}$, $Z_0 \leq 7\omega^{1/4}M$ and $M/4 < M_{t-1} < M$, we obtain

$$\mathbb{E}[d(w_t)-2] \leq \frac{t-1}{M} + 4(1+9\omega^{1/5})(14\omega^{1/4}) + 18\omega^{1/5} \leq -\frac{t}{M} + 19\omega^{1/5},$$

Proving our induction hypothesis.

7. Conclusion and open problems

In this document, we have followed the existence of the discrete duality principle through several settings. Starting from the branching process approach (Theorem 2.7), with the specific case of the Poisson branching process and Theorem 2.10. Along with giant component studies, we have then seen its first and most famous application in graph theory with the Erdős-Rényi model (Theorem 3.11). Then we have explored a first generalisation to some degree sequences under technical conditions, Theorem 5.6. Finally we began studying more general settings with the switching methodology.

As we have seen in the switching section 6, the question of the giant component's existence is solved for any degree sequence. However, some questions solved in the Erdős-Rényi model or in the configuration model, remain open when applied to general degree sequences:

(1) In the supercritical regime, what is the asymptotic size of the giant component? In the Erdős-Rényi model, we know that $L_1 \sim \zeta n$ with ζ the smallest solution to $1 - \zeta = e^{-c\zeta}$, and in the configuration model ζ is derived from the equation $\chi(\alpha) = 0$. We believe that the use of branching processes (as per [BR15] in the configuration model or as per Erdős-Rényi study) can lead to some better results in the general setting, while the differential equation methodology seems unlikely to work in such unrestricted setting.

(2) In the supercritical regime, is the giant component unique? This is the case in the Erdős-Rényi model as well as in the configuration model. However, the general setting imposing no condition on $\mathcal{D}$, this might behave differently. Even if it seems unlikely, uniqueness could depend on additional constraints.

(3) In the subcritical regime, what is the asymptotic size of the largest component? In the proof that $G(\mathcal{D})$ has a.s. no component of linear order, there are a number of approximations in equalities involving ω and M. If we can refine them, we should be able to find a good asymptotic estimation of the largest component in the subcritical regime.

By solving these questions, we would be able to answer our main open problem : **is the discrete duality principle true for any given degree sequence?**

A. Appendices

A.1 Asymptotic notations

Notation	Name	Asymptotic Description	Limit Definition
$f(n) = o(g(n))$	Small o	f is dominated by g	$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$
$f(n) = O(g(n))$	Big o	$\lvert f\rvert$ is bounded above by g	$\limsup_{n\to\infty} \frac{\lvert f(n)\rvert}{g(n)} < \infty$
$f(n) = \Theta(g(n))$	Big theta	f is bounded above and below by g	$f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$
$f(n) \sim g(n)$	On the order of	f is equal to g	$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$
$f(n) = \Omega(g(n))$	Big omega	f is bounded below by g	$\liminf_{n\to\infty} \frac{f(n)}{g(n)} > 0$
$f(n) = \omega(g(n))$	Small omega	f dominates g	$\lim_{n\to\infty} \left\lvert \frac{f(n)}{g(n)} \right\rvert = \infty$

Table 6: Asymptotic notations

A.2 Probability distributions

Name	Notation	Parameters	Support	PDF or PMF.	Mean	Variance
Bernoulli	$Ber(p)$	$0 \leq p \leq 1$	$k \in \{0, 1\}$	$\begin{cases} 1-p & \text{if } k = 0 \\ p & \text{if } k = 1 \end{cases}$	p	$p(1-p)$
Binomial	$Bin(n, p)$	n trials, each prob. p success	$k \in [n]$	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson	$Po(\lambda)$	$\lambda > 0$ rate	$k \in \mathbb{N} \cup \{0\}$	$e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ
Normal	$\mathcal{N}(\mu, \sigma^2)$	$\mu \in \mathbb{R}$, $\sigma^2 > 0$	$x \in \mathbb{R}$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

Table 7: Distributions used in this document

A.3 Bounds

In this section we summarise and prove a number of useful probabilistic bounds.

Proposition A.1 (Markov inequality)**.** *Let X be a non-negative random variable, with $\mathbb{E}[X] < \infty$. Then*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

Proof.

$$\mathbb{E}[X] \geq \mathbb{E}\left[X\mathbb{1}_{\{X\geq A\}}\right] \geq a\mathbb{P}(X \geq a) \Longrightarrow \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

□

Proposition A.2 (Chebychev inequality)**.** *Let X be an integer valued random variable with $\mathbb{V}ar(X) = \sigma^2$. Then*

$$\mathbb{P}\left(|X - \mathbb{E}[X]| \geq a\right) \leq \frac{\sigma^2}{a^2}$$

In particular, with $\mathbb{E}[X] \geq \mu$,

$$\mathbb{P}(X = 0) \leq \frac{\sigma^2}{\mu^2}$$

Proof. Note that, with $Y = (X - \mathbb{E}[X])^2$

$$\mathbb{P}\left(|X - \mathbb{E}[X]| \geq a\right) = \mathbb{P}[Y \geq a^2]$$

Therefore applying Markov's inequality to Y:

$$\mathbb{P}\left(|X - \mathbb{E}[X]| \geq a\right) \leq \frac{\mathbb{E}[Y]}{a^2} = \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2} = \frac{\mathbb{V}ar(X)}{a^2}$$

For the second inequality note that

$$\mathbb{P}[X = 0] \leq \mathbb{P}\big[|X - \mathbb{E}[X]| \geq \mathbb{E}[X]\big] \leq \frac{\mathbb{V}ar(X)}{\mathbb{E}[X]^2}$$

□

Proposition A.3 (Chernoff bounds)**.** *Let $\{X_i\}$ be a sequence of i.i.d random variables. Then,*

(1) For $a \geq \mathbb{E}[X_1]$,

$$P\left[\sum_{i=1}^{n} X_i \geq na\right] \leq e^{-n\mathcal{I}(a)}$$

With $\mathcal{I}(a) = \sup_{t\geq 0}\left(ta - \log \mathbb{E}[e^{tX_1}]\right)$

(2) For $a \leq \mathbb{E}[X_1]$,

$$P\left[\sum_{i=1}^{n} X_i \leq na\right] \leq e^{-n\mathcal{I}(a)}$$

With $\mathcal{I}(a) = \sup_{t\leq 0}\left(ta - \log \mathbb{E}[e^{tX_1}]\right)$

Proof. The generic Chernoff bound for a random variable X is attained by applying Markov's inequality to e^{tX}. We are proving here the case $a \geq \mathbb{E}[X_1]$. For every $t > 0$:

$$\mathbb{P}[X \geq na] = \mathbb{P}\big[e^{tX} \geq e^{tna}\big] \leq \frac{\mathbb{E}[e^{tX}]}{e^{tna}}$$

When X is the sum of n random variables $X_1, ..., X_n$, we get for any $t > 0$,

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \geq na\right] \leq e^{-tna}\mathbb{E}\left[\prod_{i=1}^{n} e^{tX_i}\right]$$

In particular, optimizing over t and using the assumption that X_i are independent, we obtain,

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \geq na\right] \leq \min_{t\geq 0} e^{-tna}\prod_{i=1}^{n}\mathbb{E}\left[e^{tX_i}\right]$$

Assuming that the X_i are i.i.d we obtain,

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \geq na\right] \leq \min_{t\geq 0} e^{-tna}\left(\mathbb{E}\left[e^{tX_1}\right]\right)^n = e^{-n\sup_{t\geq 0}\left(ta-\log\mathbb{E}\left[e^{tX_1}\right]\right)} = e^{-n\mathcal{I}(a)}$$

□

Remark A.4. Let $X_1, \dots, X_N$ be a set of independent Bernoulli random variables with expected value p and let $X = \sum_{i=1}^{n} X_i$. Then for every $0 < t < Np$,

$$\mathbb{P}\left[|X - \mathbb{E}[X]| > t\right] \leq 2e^{-\frac{t^2}{3Np}}$$

Proposition A.5 (Azuma's inequality)**.** *Let X be a random variable determined by a sequence of N random experiments $T_1, \dots, T_n$ such that for every $1 \leq i \leq N$ and for any possible sequences $t_1, \dots, t_{i-1}, t_i$ and $t_1, \dots, t_{i-1}, t_i'$ such that*

$$\left|\mathbb{E}[X \mid T_1 = t_1, \dots, T_i = t_i] - \mathbb{E}[X \mid T_1 = t_1, \dots, T_i = t_i']\right| \leq c_i$$

for some $c_i > 0$. Then

$$\mathbb{P}\left[|X - \mathbb{E}[X]| > t\right] < 2e^{-\frac{t^2}{2\sum_{i=1}^{N} c_i^2}}$$

This yields the following standard corollary

Corollary A.6. *Let $\Sigma = \Sigma_1, \Sigma_2, \dots, \Sigma_n$ be a sequence of random events. Let $f(\Sigma)$ be a random variable defined by these Σ_i. If for each i,*

$$\max\left|\mathbb{E}\left[f(\Sigma) \mid \Sigma_1, \Sigma_2, \dots, \Sigma_{i+1}\right] - \mathbb{E}\left[f(\Sigma) \mid \Sigma_1, \Sigma_2, \dots, \Sigma_i\right]\right| \leq c_i,$$

then

$$\mathbb{P}\left[|f - \mathbb{E}[f]| > t\right] \leq 2e^{\frac{-t^2}{2\sum c_i^2}}.$$

A.4 Erdős-Gallai theorem

As stated in 4, all degree sequences are not feasible. The main issue arise if the sequence includes too many high degree vertices. The configuration model avoids this issue because it generates only sparse graphs, hence with no high degree vertices.

In more general situation, there exists some results solving the feasibility of degree sequences. One of them is the Erdős–Gallai theorem. It gives a necessary and sufficient condition for a finite sequence of natural numbers to be the degree sequence of a simple graph.

Theorem A.7. *A sequence of non-negative integers $d_1 \geq ... \geq d_n$ can be represented as the degree sequence of a finite simple graph on n vertices if and only if $d_1 + ... + d_n$ is even and:*

$$\sum_{i=1}^{k} d_i \leq k(k-1) + \sum_{i=k+1}^{n} \min(d_i, k)$$

holds for every k in $1 \leq k \leq n$.

It is not difficult to show that the conditions of the Erdős-Rényi theorem are necessary for a sequence of numbers to be graphic. The first part relate to hand-shaking lemmas. The second part can be proven by double counting argument. The original proof that these conditions are sufficient was given by Erdos and Gallai in 1960 and was rather lengthy. In 1986 Choudum gave a simpler proof (see [Cho86]) working by induction on $s = d_1 + ... + d_n$.

B. Bibliography

References

[ACL00] William Aiello, Fan Chung, and Linyuan Lu. A Random Graph Model for Massive Graphs. In *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing*, STOC '00, pages 171–180, New York, NY, USA, 2000. ACM. event-place: Portland, Oregon, USA.

[AS92] Noga Alon and Joel H Spencer. *The Probabilistic Method*. Wiley edition, 1992.

[BC78] Edward A Bender and E. Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, May 1978.

[Bol80] Béla Bollobás. A Probabilistic Proof of an Asymptotic Formula for the Number of Labelled Regular Graphs. *European Journal of Combinatorics*, 1(4):311–316, December 1980.

[BR15] Bela Bollobas and Oliver Riordan. An old approach to the giant component problem. *Journal of Combinatorial Theory, Series B*, 113:236–260, July 2015. arXiv: 1209.3691.

[Cho86] S. A. Choudum. A simple proof of the Erdos-Gallai theorem on graph sequences. *Bulletin of the Australian Mathematical Society*, 33(1):67–70, February 1986.

[Chv91] V. Chvátal. Almost all graphs with 1.44n edges are 3-colorable. *Random Structures & Algorithms*, 2(1):11–28, 1991.

[Fel] William Feller. An Introduction to Probability Theory and Its Applications.

[FFF99] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos. On Power-law Relationships of the Internet Topology. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, SIGCOMM '99, pages 251–262, New York, NY, USA, 1999. ACM. event-place: Cambridge, Massachusetts, USA.

[Hof16] Remco van der Hofstad. *Random Graphs and Complex Networks*, volume 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2016.

[JL07] Svante Janson and Malwina Luczak. A new approach to the giant component problem. *arXiv:0707.1786 [math]*, July 2007. arXiv: 0707.1786.

[JLR11] Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random Graphs*. John Wiley & Sons, September 2011. Google-Books-ID: RjnqVoB4VmUC.

[JPRR18] Felix Joos, Guillem Perarnau, Dieter Rautenbach, and Bruce Reed. How to determine if a random graph with a fixed degree sequence has a giant component. *Probability Theory and Related Fields*, 170(1-2):263–310, February 2018. arXiv: 1601.03714.

[KS08] M. Kang and T. G. Seierstad. The Critical Phase for Random Graphs with a Given Degree Sequence. *Combinatorics, Probability and Computing*, 17(1):67–86, January 2008.

[LP17] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. American Mathematical Soc., October 2017. Google-Books-ID: f208DwAAQBAJ.

[McK85] Brendan D. McKay. Asymptotics for Symmetric 0-1 Matrices with Prescribed Row Sums. *Ars. Combinatorica*, 1985.

[McK11] Brendan D. McKay. Subgraphs of Random Graphs with Specified Degrees. In *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010)*, pages 2489–2501. Published by Hindustan Book Agency (HBA), India. WSPC Distribute for All Markets Except in India, June 2011.

[MR95] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, 6(2-3):161–180, 1995.

[MR98] Michael Molloy and Bruce Reed. The Size of the Giant Component of a Random Graph with a Given Degree Sequence. *Combinatorics, Probability and Computing*, 7(3):295–305, September 1998.

[Pet91] Julius Petersen. Theory of regular graphs, 1891.

[Pit08] Boris Pittel. Edge percolation on a random regular graph of low degree. *The Annals of Probability*, 36(4):1359–1389, July 2008. arXiv: 0808.3516.

[Rio12] Oliver Riordan. The phase transition in the configuration model. *Combinatorics, Probability and Computing*, 21(1-2):265–299, March 2012. arXiv: 1104.0613.

[Łu90] Tomasz Łuczak. Component behavior near the critical point of the random graph process. *Random Structures & Algorithms*, 1(3):287–310, 1990.

[Wor80] Nicholas C. Wormald. Some problems in the enumeration of labelled graphs. *Bulletin of the Australian Mathematical Society*, 21(1):159–160, February 1980.

[Wor95] Nicholas C. Wormald. Differential Equations for Random Processes and Random Graphs. *The Annals of Applied Probability*, 5(4):1217–1235, 1995.

[Wor99] N. C. Wormald. Models of Random Regular Graphs, July 1999.